

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Comment comprendre le rôle des mathématiques dans l'enseignement secondaire à travers les prévisions météorologiques ?

MOSSOUX, Margaux

Award date:
2015

Awarding institution:
Université de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**UNIVERSITÉ
DE NAMUR**

FACULTÉ
DES SCIENCES

UNIVERSITÉ DE NAMUR

Faculté des Sciences

**COMMENT COMPRENDRE LE RÔLE DES MATHÉMATIQUES DANS
L'ENSEIGNEMENT SECONDAIRE À TRAVERS LES PRÉVISIONS
MÉTÉOROLOGIQUES ?**

**Mémoire présenté pour l'obtention du grade académique de Master en
Sciences mathématiques à finalité didactique**

Margaux MOSSOUX

Août 2015



UNIVERSITÉ DE NAMUR
Faculté des Sciences

**COMMENT COMPRENDRE LE RÔLE DES MATHÉMATIQUES DANS
L'ENSEIGNEMENT SECONDAIRE À TRAVERS LES PRÉVISIONS
MÉTÉOROLOGIQUES ?**

Promoteur :

Annick Sartenaer

**Mémoire présenté pour l'obtention du grade académique de Master en
Sciences mathématiques à finalité didactique**

Margaux MOSSOUX

Août 2015

Résumé

Ce mémoire traite du rôle des mathématiques enseignées au degré supérieur des études secondaires dans le monde actuel, le but étant de faire comprendre au plus grand nombre ce rôle à travers les prévisions météorologiques.

Nous introduisons les prévisions météorologiques en retraçant l'histoire, en étudiant l'évolution en termes de performances et efficacité. Nous étudions ensuite le programme de mathématique au dernier cycle du secondaire, afin d'en tirer les informations en lien avec le calcul numérique. Nous développons la théorie du calcul numérique et illustrons par quelques algorithmes généralisant les notions vues en secondaire. Finalement, nous expliquons le principe d'assimilation de données dans les prévisions numériques du temps, qui constitue une illustration de résolution de problèmes mathématiques.

Mots-clés : enseignement secondaire, calcul numérique, prévisions météorologiques, assimilation de données.

Abstract

This master's thesis discusses the role of mathematics taught during the end of high school nowadays in order for a majority to understand that role via weather forecasting.

Firstly, we introduce the history of weather forecasting after which we study its evolution in terms of performance and efficiency. Then we study the mathematical cursus program for the last cycle of high school and isolate the part concerning numerical calculation. We then further develop numerical calculation and illustrate it through several algorithms using notions learned during high school. Finally we explain the data assimilation principle used for the numerical weather forecasting as an illustration of mathematical problem resolution.

Keywords : high school, numerical calculation, weather forecasting, data assimilation.

Remerciements

*Je remercie toutes les personnes qui ont, de près ou de loin,
participé à l'élaboration de ce mémoire.*

*En particulier, je remercie ma promotrice Annick Sartenaer pour sa
patience, pour ses recherches et indications. Je la remercie de m'avoir
toujours encouragée et aiguillée dans mon travail.*

*Merci à elle d'avoir accepté de m'accompagner dans ce sujet que nous
avons imaginé et fait évoluer ensemble.*

*Je tiens également à remercier ma famille et mes amis, pour la relecture
de ce mémoire ainsi que pour leur soutien moral non seulement durant la
réalisation de celui-ci, mais aussi tout au long de mes études.*

Table des matières

Introduction	5
1 Les prévisions météorologiques au cours du temps	6
2 Les mathématiques dans l'enseignement secondaire	9
2.1 Programme scolaire	9
2.2 Matrices	11
2.3 Systèmes d'équations linéaires	21
2.4 Dérivées et équations aux dérivées partielles	24
2.5 Problèmes d'optimisation sans contrainte	27
2.6 Statistiques	28
3 Le calcul numérique	32
3.1 Arithmétique en virgule flottante	32
3.2 Opérations et erreurs d'arrondi	36
3.3 Algorithmes	37
3.4 Illustrations	39
4 Le principe de prévision numérique du temps	49
4.1 Discrétisation de l'atmosphère	49
4.2 Observations météorologiques	51
4.3 Modélisation numérique	52
4.4 Assimilation de données variationnelle	52
4.5 Liens avec les notions vues en secondaire	56
Conclusion	58
Bibliographie	59

Introduction

La problématique de l'intérêt que les jeunes portent à leur scolarité, et plus particulièrement aux mathématiques, semble être une piste intéressante, voire évidente, pour le choix du sujet de ce mémoire, d'autant plus pour une future enseignante. Le but ultime de ce travail est ainsi de montrer l'intérêt des mathématiques dans la société actuelle avec, en ligne de mire, les prévisions météorologiques. Dès lors, nous proposons d'illustrer la résolution de problèmes mathématiques via le calcul numérique et, plus particulièrement, l'assimilation de données, qui consiste à intégrer des observations d'un système afin d'en ajuster l'état obtenu numériquement. Ces illustrations sont à destination d'élèves du secondaire - degré supérieur.

Dans un premier temps, nous retraçons l'historique des prévisions météorologique dans le chapitre 1. Nous y voyons l'évolution en termes de performance des ordinateurs et de précision des prévisions.

Dans le chapitre 2, nous analysons les matières que les élèves du degré supérieur abordent en mathématique, en lien avec le calcul numérique. Il s'agit donc ici de repérer les matières pour lesquelles une résolution « manuelle » serait trop compliquée et dans lesquelles la calculatrice graphique, ou autre logiciel de résolution numérique, est utilisée. Nous donnons différents exemples afin de comprendre les notions développées.

Ensuite, dans le chapitre 3, nous généralisons certaines des matières retenues en vue de l'utilisation d'un ordinateur. Nous analysons alors les possibilités et les limites qu'offrent les ordinateurs et en étudions brièvement le principe de stockage. Nous donnons quelques exemples afin d'illustrer comment les mathématiciens généralisent les méthodes de résolution de problèmes tels que vus en secondaire afin d'aboutir à un algorithme.

Nous développons finalement une application concrète au chapitre 4 : les prévisions météorologiques. Remarquons que l'objectif de ce mémoire n'est pas de détailler chaque point des techniques de prévisions météorologiques mais d'en avoir une vision d'ensemble afin de comprendre le principe général et de prendre conscience du rôle des mathématiques.

Chapitre 1

Les prévisions météorologiques au cours du temps

Les informations ci-dessous proviennent des articles de COIFFIER J. [19] et de EC JOURNAL [20].

Prévoir le temps qu'il fera plus tard semble aujourd'hui être quelque chose de très anodin. Il suffit de prendre son smartphone, de se connecter à internet, d'ouvrir un journal ou encore de regarder la télévision pour obtenir les informations concernant la température, la vitesse du vent, les précipitations, etc. Et pourtant, cette démarche de prévision météorologique était à peine imaginable il y a quelques dizaines d'années. C'est en effet en 1904 que les prévisions météorologiques ont pris tout leur sens, lorsque le mathématicien, physicien et géophysicien norvégien Vilhelm Bjerknes déclara :

« Si, comme le pense chaque homme raisonnant scientifiquement, les phénomènes atmosphériques se développent à partir de ceux qui les précèdent suivant des lois précises, on peut en déduire que les conditions nécessaires et suffisantes pour une solution rationnelle de la prévision en météorologie sont les suivantes :

- on doit connaître avec une précision suffisante l'état de l'atmosphère à un instant donné ;*
- on doit connaître avec une précision suffisante les lois selon lesquelles un état de l'atmosphère se développe à partir de l'état précédent. »*

C'est, en quelque sorte, ce type de raisonnement qui est à la tête de l'assimilation de données, que nous verrons plus en détail dans le chapitre 4. Vilhelm Bjerknes remarqua cependant la difficulté à l'époque de résoudre ce type de problème puisqu'il s'agissait de résoudre des équations non linéaires - équations aux dérivées partielles - dont nous reparlerons dans le chapitre 2.

En 1922, le mathématicien Lewis Fry Richardson imagina une usine constituée de pas moins de 64000 « calculateurs » humains qui résoudraient simultanément des systèmes d'équations aux dérivées partielles. Ce nombre de personnes est, dit-il, le nombre nécessaire pour pouvoir faire des prévisions à échéance courte et avant que cette échéance ne se soit écoulée.

Peu après la Deuxième Guerre mondiale, en 1946 plus précisément, le mathématicien américain John von Neumann créa le premier supercalculateur au monde, appelé *Eniac* (*Electronic Numerical Integrator Analyser and Computer*) permettant de résoudre

des problèmes mathématiques de manière numérique, c'est-à-dire moyennant l'utilisation d'ordinateurs.

C'est finalement en 1950 que les premières prévisions numériques du temps ont pu voir le jour, grâce à John von Neumann mais aussi à Jules Charney et Ragnar Fjörtoft. Dès ce moment, il fallut apprendre à développer des méthodes de calculs aussi efficaces que rapides. En parallèle, les ordinateurs ont évolué de manière spectaculaire. Si l'on regarde la figure 1.1, on constate que l'*Eniac* pouvait, par exemple, effectuer près de 1000 opérations par seconde (appelées *flops*, voir chapitre 3) à l'époque, alors que, de nos jours, le *Cray XT5* est capable d'en effectuer plus de 10^{15} .

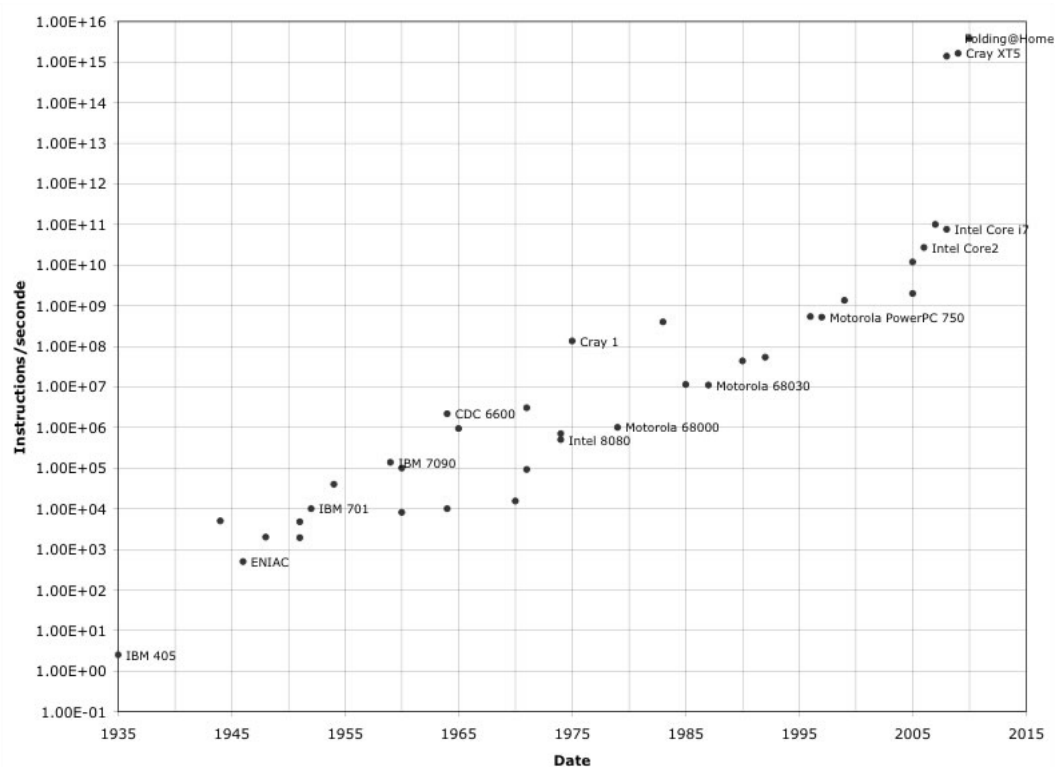


FIGURE 1.1 – Représentation de l'évolution du nombre de flops par seconde au cours du temps. Source : A. Bouquet [18].

Notons que la rapidité du calcul d'un ordinateur dépend non seulement du nombre de flops qu'il peut effectuer chaque seconde mais également des caractéristiques du modèle utilisé afin d'effectuer ledit calcul.

Au cours du temps, les prévisions météorologiques ont pu évoluer de manière à rendre plus certains les résultats. En effet, parallèlement à l'évolution des ordinateurs, les modèles mis en place ont pu également être améliorés et affinés de sorte que les prévisions peuvent, au fur et à mesure, s'étendre à une échéance plus large. On constate sur la figure 1.2 que, par exemple, une prévision à cinq jours d'échéance en 2010 est aussi fiable qu'une prévision à trois jours en 1985 dans le même hémisphère. Par ailleurs, nous notons également une diminution d'écart entre la qualité des prévisions des deux hémisphères.

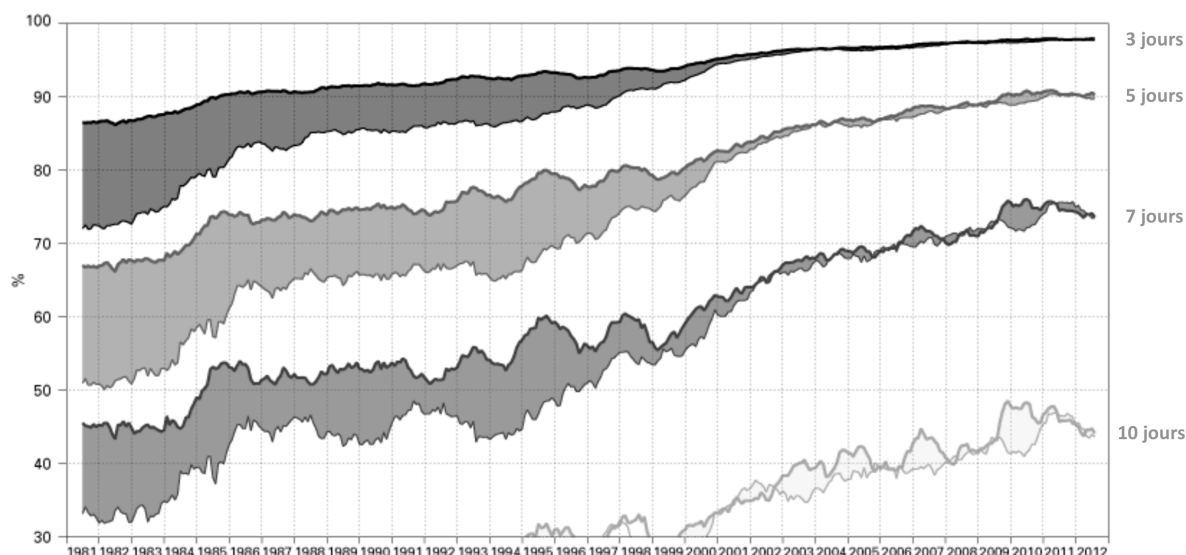


FIGURE 1.2 – Représentation de l'évolution de la qualité des prévisions météorologiques basée sur le coefficient de corrélation d'anomalie au niveau 500 hPa de l'atmosphère. Trait supérieur : hémisphère Nord ; trait inférieur : hémisphère Sud. Source : E. Anderson [16].

En 1961, le mathématicien Edward Lorenz découvrit un fait à l'époque très surprenant. En modifiant de quelques décimales les données initiales d'un problème à résoudre sur ordinateur, les solutions obtenues étaient complètement différentes. En réalité, celles-ci paraissaient très similaires au début de l'exécution du programme mais plus le temps passait et plus le comportement divergeait. Ce phénomène est encore bien connu aujourd'hui sous le nom d'« effet papillon » : Une petite perturbation des conditions initiales d'un système peut engendrer des résultats qui varient grandement. On appelle cela un phénomène chaotique. Dès lors, bien que les prévisions météorologiques soient en constante évolution, il serait utopique de penser que celles-ci atteindront un jour la perfection. De plus, nous pouvons effectuer des prévisions à court terme de manière précise mais nous ne parviendrons jamais à réaliser des prévisions dites « haute prévision » à grande échéance.

Dans le chapitre suivant, nous démarrons notre recherche par une étude du programme de mathématique dans l'enseignement secondaire. Il s'agit du point de départ de ce mémoire et nous y détaillons quelques notions utiles dans les prévisions météorologiques.

Chapitre 2

Les mathématiques dans l'enseignement secondaire

Les mathématiques prennent une place importante dans l'enseignement secondaire général. De la première année jusqu'à la quatrième, le cours de mathématique représente à lui seul cinq heures par semaine, soit entre 15% et 18% du nombre total d'heures par semaine. En cinquième et sixième année, les élèves suivent un cours de deux, quatre ou six heures, ce qui correspond à une proportion de 6% à 21%. Ces proportions dépendent des établissements scolaires ainsi que des options des élèves.

2.1 Programme scolaire

Nous nous intéressons ici à l'enseignement du troisième cycle, autrement dit des cinquième et sixième années, et plus particulièrement aux cours de mathématique de six heures par semaine. Les programmes divergent légèrement dans l'enseignement officiel (anciennement public), dépendant de la Communauté Française, et dans l'enseignement libre (anciennement privé). Nous allons ici effectuer un récapitulatif des différentes matières abordées au cours de mathématique, sous forme de listes de compétences à atteindre en fin d'année et pour l'ensemble des programmes. Les informations reprises ci-dessous proviennent des documents du MINISTÈRE DE LA COMMUNAUTÉ FRANÇAISE [25] et de la FÉDÉRATION DE L'ENSEIGNEMENT SECONDAIRE CATHOLIQUE [23]. Pour des raisons de clarté et de lisibilité, certains éléments ont volontairement été retirés des listes. Ces dernières ne sont ainsi pas exhaustives et ne représentent qu'un résumé du programme.

Dans chacun des programmes, il existe différents liens avec le calcul numérique, notamment via l'utilisation de calculatrices ou de logiciels. Ces liens sont mis en évidence dans les listes ci-dessous.

Cinquième année

Algèbre

- modéliser une situation, un ensemble de contraintes, par un tableau de nombres, appelé matrice ;
- effectuer des calculs matriciels ;
- résoudre et discuter la résolution de systèmes de trois équations à trois inconnues ;

- résoudre des systèmes d'équations linéaires et distinguer les systèmes ayant une solution, une infinité de solutions ou aucune solution ;
- étudier les propriétés du calcul matriciel pour en arriver à la détermination de la matrice inverse. « *Ce type de calcul peut faire l'objet d'un travail de programmation avec un logiciel ou une calculatrice.* » [23].

Analyse

- modéliser une situation par une suite arithmétique ou géométrique, être capable d'en déterminer le terme ;
- maîtriser les notions d'infini et de limite, adapter ces notions aux fonctions ;
- déterminer l'équation des asymptotes au graphique d'une fonction, prédire l'existence d'asymptotes d'une fonction à partir de son graphique ou de son équation ;
- être capable de lever une indétermination ;
- calculer la dérivée d'une fonction, interpréter (géométriquement et/ou physiquement) la dérivée d'une fonction en un point, utiliser les propriétés des dérivées dans diverses applications (« *modélisation de problèmes, approximation locale d'une fonction par une fonction du premier degré, [...] problèmes d'optimisation.* ») [25] ;
- étendre les manipulations sur les fonctions du second degré aux fonctions homographiques, étudier les décompositions de fonction et leur dérivation. « *L'enseignement de l'analyse s'appuie sur l'utilisation de la calculatrice graphique ou des logiciels appropriés.* » [23].

Géométrie

- effectuer et interpréter géométriquement des opérations sur des vecteurs ;
- comprendre les notions de parallélisme, d'incidence et d'homothétie ;
- construire un point de percée et une parallèle à deux droites gauches ;
- interpréter la géométrie vectorielle en termes de systèmes d'équations linéaires.

Trigonométrie

- utiliser les formules trigonométriques ;
- résoudre des équations et inéquations trigonométriques, représenter les solutions sur le cercle trigonométrique. « *Quelques équations [...] ; seront résolues de manière approchée, par itération (utiliser une calculatrice graphique ou un logiciel).* » [23].

Sixième année

Analyse

- connaître les définitions et opérations concernant les nombres complexes, les représenter géométriquement ;
- convertir les nombres complexes de la forme algébrique à la forme trigonométrique et inversement ;
- faire une étude des fonctions cyclométriques, dériver des fonctions cyclométriques ;

- comprendre les notions de primitives et d'intégrales ;
- faire une étude des fonctions logarithmiques et exponentielles.

Géométrie

- faire une étude des coniques.

Statistiques

- déterminer une droite de régression (**méthode des moindres carrés**) et un coefficient de corrélation ;
- comprendre la notion d'ajustement linéaire d'un nuage de points (**méthode des moindres carrés**) ;
- comprendre les différentes notions de l'analyse combinatoire et des probabilités.

Pour la suite du travail, nous nous intéressons aux branches dans lesquelles nous avons pu observer des liens avec le calcul numérique :

- matrices ;
- problèmes d'optimisation ;
- dérivée d'une fonction ;
- approximation par méthode itérative ;
- méthode des moindres carrés.

2.2 Matrices

Avant toute chose, commençons par rappeler quelques notions théoriques sur les matrices. Nous étudions également quelques exemples afin de comprendre ces différentes notions. Les définitions et remarques de cette section sont inspirées des livres de LAY D. C. [13] et de QUARTERONI A., SACCO R. et SALERI F. [15]. Elles ont été adaptées de telle sorte qu'elles soient compréhensibles par des élèves du secondaire. Par ailleurs, nous ne considérons que l'espace des nombres réels.

Définition 2.1. Soient deux entiers m et n . Une *matrice* $m \times n$ est un ensemble de mn nombres réels¹ a_{ij} , où $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$. On représente typiquement une matrice de la manière suivante :

$$A_{(m \times n)} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

1. Une matrice est en réalité un ensemble de nombres scalaires. Un nombre scalaire peut être soit réel, soit complexe. Dans ce mémoire, pour une question de facilité de compréhension, nous ne considérons que le cas réel. Notons aussi que nous nous tenons, dans les exemples, à des valeurs entières par souci de clarté mais ceci n'est, bien entendu, pas restrictif.

Les valeurs i et j sont appelées *indices de ligne et de colonne*, respectivement. On dit qu'une matrice A est *carrée* si $m = n$. Celle-ci est alors dite d'ordre n et on la note A_n . La *diagonale* d'une matrice A , carrée d'ordre n , est l'ensemble des éléments :

$$\{a_{11}, a_{22}, \dots, a_{nn}\} .$$

Une matrice carrée est dite *triangulaire inférieure* (respectivement *supérieure*) si tous les éléments situés au-dessus (respectivement en dessous) de la diagonale de cette matrice sont nuls.

Exemple 2.1. La matrice :

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 3 & 4 \\ 0 & 0 & 7 \end{pmatrix}$$

est une matrice carrée d'ordre 3. Elle est triangulaire supérieure.

Remarque. Lorsque $i = 1$ (respectivement $j = 1$), la matrice est appelée *vecteur ligne* (respectivement *colonne*). Dans ce cas, on note $x \in \mathbb{R}^n$ plutôt que $X_{(1 \times n)}$ (respectivement $X_{(n \times 1)}$). Par convention, on considère qu'un vecteur de \mathbb{R}^n est un vecteur colonne.

La somme matricielle est relativement analogue à la somme de deux scalaires. Elle s'effectue très simplement. Il en est de même pour la multiplication d'une matrice par un scalaire. Le produit matriciel est, quant à lui, plus compliqué.

Définition 2.2. La *somme* de deux matrices $A_{(m \times n)} = (a_{ij})$ et $B_{(m \times n)} = (b_{ij})$ est la matrice $C_{(m \times n)} = (c_{ij})$ définie par :

$$c_{ij} = a_{ij} + b_{ij} ,$$

où $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$.

La *matrice nulle* est définie par :

$$0_{(m \times n)} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

et est telle que le résultat de l'addition d'une matrice quelconque avec celle-ci vaut la matrice de départ. On dit que la matrice nulle est *neutre pour l'addition*.

Exemple 2.2. Soient les matrices :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 3 \end{pmatrix} .$$

Nous avons $m = 2$ et $n = 3$. Les matrices A et B étant de même dimension, nous pouvons calculer la somme de ces deux matrices :

$$\begin{aligned} A + B &= \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1+2 & 2+3 & 4+5 \\ 3+1 & 1+2 & 1+3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 5 & 9 \\ 4 & 3 & 4 \end{pmatrix} . \end{aligned}$$

Définition 2.3. La *multiplication* de la matrice $A_{(m \times n)} = (a_{ij})$ par le nombre réel α est la matrice $C_{(m \times n)} = (c_{ij})$ définie par :

$$c_{ij} = \alpha \cdot a_{ij} ,$$

où $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$.

Exemple 2.3. Soit la matrice :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \end{pmatrix} .$$

Nous souhaitons multiplier cette matrice par le nombre réel 2. Nous obtenons ainsi

$$\begin{aligned} 2 \cdot A &= 2 \cdot \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \cdot 1 & 2 \cdot 2 & 2 \cdot 4 \\ 2 \cdot 3 & 2 \cdot 1 & 2 \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 4 & 8 \\ 6 & 2 & 2 \end{pmatrix} . \end{aligned}$$

Définition 2.4. Le *produit* de deux matrices $A_{(m \times p)} = (a_{ij})$ et $B_{(p \times n)} = (b_{ij})$ est la matrice $C_{(m \times n)} = (c_{ij})$ définie par :

$$c_{ij} = \sum_{k=1}^p a_{ik} \cdot b_{kj} ,$$

où $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$.

La *matrice identité d'ordre n* est une matrice carrée définie par :

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

et est telle que le résultat de la multiplication de celle-ci par une autre matrice carrée à n lignes vaut cette matrice. De même, le résultat de la multiplication d'une matrice à n colonnes par la matrice identité vaut la matrice de départ. On dit que la matrice identité est *neutre pour la multiplication*.

Remarque. Même si le produit d'une matrice A par une matrice B existe, il se peut que le produit de la matrice B par la matrice A n'existe pas. Pour qu'il existe, il faut que les dimensions de chacune des deux matrices soient identiques.

Exemple 2.4. Soient les matrices :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix} .$$

Nous avons $m = 2$, $p = 3$ et $n = 3$. Le nombre de colonnes dans la matrice A étant identique au nombre de lignes dans la matrice B , nous pouvons calculer le produit de ces deux matrices :

$$\begin{aligned} A \cdot B &= \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 4 + 2 \cdot 2 + 4 \cdot 3 & 1 \cdot 5 + 2 \cdot 1 + 4 \cdot 1 & 1 \cdot 2 + 2 \cdot 1 + 4 \cdot 4 \\ 3 \cdot 4 + 1 \cdot 2 + 1 \cdot 3 & 3 \cdot 5 + 1 \cdot 1 + 1 \cdot 1 & 3 \cdot 2 + 1 \cdot 1 + 1 \cdot 4 \end{pmatrix} \\ &= \begin{pmatrix} 4 + 4 + 12 & 5 + 2 + 4 & 2 + 2 + 16 \\ 12 + 2 + 3 & 15 + 1 + 1 & 6 + 1 + 4 \end{pmatrix} \\ &= \begin{pmatrix} 20 & 11 & 20 \\ 17 & 17 & 11 \end{pmatrix} . \end{aligned}$$

Cependant, le produit $B \cdot A$ n'existe pas puisque le nombre de colonnes dans la matrice B est différent du nombre de lignes dans la matrice A .

Remarque. Le produit matriciel n'est pas commutatif. En d'autres termes, le produit d'une matrice A par une matrice B n'est pas nécessairement égal au produit de la matrice B par la matrice A . L'exemple suivant le prouve.

Exemple 2.5. Soient les matrices :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \\ 5 & 4 & 3 \end{pmatrix} \text{ et } B = \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix}.$$

Nous avons, d'une part, que le produit de la matrice A par la matrice B vaut :

$$A \cdot B = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \\ 5 & 4 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 20 & 11 & 20 \\ 17 & 17 & 11 \\ 37 & 32 & 26 \end{pmatrix}$$

et, d'autre part, que le produit de la matrice B par la matrice A vaut :

$$B \cdot A = \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \\ 5 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 29 & 21 & 27 \\ 10 & 9 & 12 \\ 26 & 23 & 25 \end{pmatrix}.$$

Nous constatons effectivement que $A \cdot B \neq B \cdot A$.

Définition 2.5. La *transposée* d'une matrice $A_{(m \times n)} = (a_{ij})$ est une matrice $n \times m$ définie par :

$$A^T = (a_{ji}) = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

On dit que la matrice A est *symétrique* si $A = A^T$. La matrice A est dite *anti-symétrique* si $A = -A^T$.

Remarque. La transposée du produit d'une matrice A par une matrice B est égale au produit de la transposée de la matrice B par la transposée de la matrice A . L'exemple suivant illustre cette propriété.

Exemple 2.6. Soient les matrices :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \\ 5 & 4 & 3 \end{pmatrix} \text{ et } B = \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix}.$$

Nous avons, d'une part, que la transposée du produit de la matrice A par la matrice B vaut :

$$(A \cdot B)^T = \begin{pmatrix} 20 & 11 & 20 \\ 17 & 17 & 11 \\ 37 & 32 & 26 \end{pmatrix}^T = \begin{pmatrix} 20 & 17 & 37 \\ 11 & 17 & 32 \\ 20 & 11 & 26 \end{pmatrix}$$

et, d'autre part, que le produit de la transposée de la matrice B par la transposée de la matrice A vaut :

$$\begin{aligned} B^T \cdot A^T &= \begin{pmatrix} 4 & 5 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 1 \\ 5 & 4 & 3 \end{pmatrix}^T \\ &= \begin{pmatrix} 4 & 2 & 3 \\ 5 & 1 & 1 \\ 2 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 5 \\ 2 & 1 & 4 \\ 4 & 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 20 & 17 & 37 \\ 11 & 17 & 32 \\ 20 & 11 & 26 \end{pmatrix}. \end{aligned}$$

Nous constatons effectivement que $(A \cdot B)^T = B^T \cdot A^T$.

Définition 2.6. Soit une matrice $A_{(m \times n)}$. On appelle *sous-matrice* de A une matrice $(k \times l)$, où $k \leq m$ et $l \leq n$, dont les éléments forment un bloc contenu dans la matrice A . On note A_{kl} une sous-matrice de A .

Exemple 2.7. Soit la matrice :

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 4 & 1 & 2 & 5 & 5 \\ 5 & 3 & 5 & 1 & 3 \\ 2 & 4 & 3 & 2 & 4 \end{pmatrix}.$$

Nous pouvons, par exemple, décomposer A en sous-matrices de la façon suivante :

$$A = \left(\begin{array}{cc|ccc} 1 & 2 & 3 & 1 & 2 \\ 4 & 1 & 2 & 5 & 5 \\ \hline 5 & 3 & 5 & 1 & 3 \\ 2 & 4 & 3 & 2 & 4 \end{array} \right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

et nous obtenons les quatre sous-matrices de A suivantes :

$$A_{11} = \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} ; A_{12} = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 5 & 5 \end{pmatrix} ; A_{21} = \begin{pmatrix} 5 & 3 \\ 2 & 4 \end{pmatrix} \text{ et } A_{22} = \begin{pmatrix} 5 & 1 & 3 \\ 3 & 2 & 4 \end{pmatrix} .$$

Nous pouvons également décomposer A en sous-matrices de la façon suivante :

$$A = \left(\begin{array}{c|ccc|c} 1 & 2 & 3 & 1 & 2 \\ \hline 4 & 1 & 2 & 5 & 5 \\ 5 & 3 & 5 & 1 & 3 \\ \hline 2 & 4 & 3 & 2 & 4 \end{array} \right) = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}$$

et nous obtenons les six sous-matrices de A suivantes :

$$A_{11} = \begin{pmatrix} 1 \end{pmatrix} ; A_{12} = \begin{pmatrix} 2 & 3 & 1 \end{pmatrix} ; A_{13} = \begin{pmatrix} 2 \end{pmatrix} ; A_{21} = \begin{pmatrix} 4 \\ 5 \\ 2 \end{pmatrix} ; A_{22} = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 5 & 1 \\ 4 & 3 & 2 \end{pmatrix} \text{ et } A_{23} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix} .$$

Définition 2.7. Soit $A = (a_{ij})$ une matrice carrée d'ordre n ($n \geq 2$). Notons $A_{\{i,j\}}$ la matrice obtenue en supprimant la ligne i et la colonne j de la matrice A . Le *déterminant* de la matrice A est alors donné par :

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{\{i,j\}}) ,$$

pour n'importe quelle valeur de i entre 1 et n . Par convention, le déterminant d'une matrice d'ordre 1, c'est-à-dire d'un scalaire, est la valeur de ce scalaire. On dit que la matrice A est *régulière*, ou *non singulière*, si $\det(A) \neq 0$.

Exemple 2.8. Soit la matrice carrée d'ordre 3 :

$$A = \begin{pmatrix} 3 & 2 & 3 \\ 4 & 1 & 4 \\ 2 & 3 & 1 \end{pmatrix} .$$

Choisissons, arbitrairement, $i = 1$. Nous obtenons ainsi

$$\begin{aligned}
 \det(A) &= \sum_{j=1}^3 (-1)^{1+j} a_{1j} \det(A_{\{1,j\}}) \\
 &= (-1)^{1+1} a_{11} \det(A_{\{1,1\}}) + (-1)^{1+2} a_{12} \det(A_{\{1,2\}}) + (-1)^{1+3} a_{13} \det(A_{\{1,3\}}) \\
 &= (-1)^2 \cdot 3 \cdot \det \begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix} + (-1)^3 \cdot 2 \cdot \det \begin{pmatrix} 4 & 4 \\ 2 & 1 \end{pmatrix} + (-1)^4 \cdot 3 \cdot \det \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \\
 &= 3 \cdot \det \begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 4 & 4 \\ 2 & 1 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \\
 &= 3 \cdot (1 \cdot 1 - 4 \cdot 3) - 2 \cdot (4 \cdot 1 - 4 \cdot 2) + 3 \cdot (4 \cdot 3 - 1 \cdot 2) \\
 &= 3 \cdot (-11) - 2 \cdot (-4) + 3 \cdot (10) \\
 &= -33 + 8 + 30 \\
 &= 5.
 \end{aligned}$$

Définition 2.8. Soit $A_{(m \times n)}$ une matrice. Le *rang* de A est l'ordre de la plus grande sous-matrice carrée non singulière de A . On dit que la matrice A est *de rang plein* si le rang de A est la plus petite valeur entre m et n .

Définition 2.9. Une matrice symétrique A d'ordre n est dite *définie positive* si, pour tout vecteur $x \in \mathbb{R}^n$ non nul, on a $x^T A x > 0$.

Une matrice symétrique A d'ordre n est dite *semi-définie positive* si, pour tout vecteur $x \in \mathbb{R}^n$ non nul, on a $x^T A x \geq 0$.

Exemple 2.9. Montrons que la matrice :

$$A = \begin{pmatrix} 2 & 1 \\ -3 & 2 \end{pmatrix}$$

est définie positive. Soit $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. On a :

$$\begin{aligned}
 x^T A x &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
 &= \begin{pmatrix} 2x_1 - 3x_2 & x_1 + 2x_2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
 &= 2x_1^2 - 3x_1x_2 + x_1x_2 + 2x_2^2 \\
 &= 2(x_1^2 + x_2^2 - x_1x_2) \\
 &> 0
 \end{aligned}$$

quels que soient x_1, x_2 , d'où A est définie positive.

La proposition ci-dessous nous donne une condition nécessaire et suffisante pour déterminer si une matrice carrée symétrique est définie positive. Il s'agit du *critère de Sylvester*.

Définition 2.10. Soit une matrice carrée d'ordre n . On appelle *mineurs principaux* de A les déterminants suivant :

$$\det(A_{\{i,i\}}) \quad \forall i = 1, \dots, n .$$

Proposition 2.1. Une matrice carrée symétrique est définie positive si et seulement ses mineurs principaux sont positifs.

Définition 2.11. Soit A une matrice carrée d'ordre n . Si A est régulière, alors la *matrice inverse* de A est la matrice notée A^{-1} telle que $A \cdot A^{-1} = A^{-1} \cdot A = I_n$. En pratique, nous pouvons calculer la matrice inverse de A par la formule :

$$A^{-1} = \frac{1}{\det(A)} C^T ,$$

où la matrice $C = (c_{ij})$, carrée d'ordre n , appelée *comatrice* de A , ou *matrice des cofacteurs* de A , est définie par

$$c_{ij} = (-1)^{i+j} \cdot \det(A_{\{i,j\}}) .$$

Remarque. L'inverse du produit d'une matrice régulière A par une matrice régulière B est égale au produit de l'inverse de la matrice B par l'inverse de la matrice A . L'exemple suivant illustre cette propriété.

Exemple 2.10. Soient les matrices :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 4 & 5 \\ 2 & 1 \end{pmatrix} .$$

Nous avons, d'une part, que l'inverse du produit de la matrice A par la matrice B vaut :

$$\begin{aligned} (A \cdot B)^{-1} &= \begin{pmatrix} 8 & 7 \\ 14 & 16 \end{pmatrix}^{-1} \\ &= \frac{1}{30} \cdot \begin{pmatrix} 16 & -14 \\ -7 & 8 \end{pmatrix}^T \\ &= \frac{1}{30} \cdot \begin{pmatrix} 16 & -7 \\ -14 & 8 \end{pmatrix} \\ &= \begin{pmatrix} \frac{16}{30} & \frac{-7}{30} \\ \frac{-14}{30} & \frac{8}{30} \end{pmatrix} \end{aligned}$$

et, d'autre part, que le produit de l'inverse de la matrice B par l'inverse de la matrice A vaut :

$$\begin{aligned}
 B^{-1} \cdot A^{-1} &= \frac{-1}{6} \cdot \begin{pmatrix} 1 & -2 \\ -5 & 4 \end{pmatrix}^T \cdot \frac{-1}{5} \cdot \begin{pmatrix} 1 & -3 \\ -2 & 1 \end{pmatrix}^T \\
 &= \frac{-1}{6} \cdot \begin{pmatrix} 1 & -5 \\ -2 & 4 \end{pmatrix} \cdot \frac{-1}{5} \cdot \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{-1}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{-4}{6} \end{pmatrix} \cdot \begin{pmatrix} \frac{-1}{5} & \frac{2}{5} \\ \frac{3}{5} & \frac{-1}{5} \end{pmatrix} \\
 &= \begin{pmatrix} \left(\frac{-1}{6} \cdot \frac{-1}{5} + \frac{5}{6} \cdot \frac{3}{5}\right) & \left(\frac{-1}{6} \cdot \frac{2}{5} + \frac{5}{6} \cdot \frac{-1}{5}\right) \\ \left(\frac{2}{6} \cdot \frac{-1}{5} + \frac{-4}{6} \cdot \frac{3}{5}\right) & \left(\frac{2}{6} \cdot \frac{2}{5} + \frac{-4}{6} \cdot \frac{-1}{5}\right) \end{pmatrix} \\
 &= \begin{pmatrix} \frac{16}{30} & \frac{-7}{30} \\ \frac{-14}{30} & \frac{8}{30} \end{pmatrix}.
 \end{aligned}$$

Nous constatons effectivement que $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

Ci-dessous, nous définissons la norme euclidienne d'un vecteur. Cette définition, bien que n'étant pas étudiée en secondaire, nous sera utile dans le chapitre 4.

Définition 2.12. La *norme euclidienne* d'un vecteur est définie de la manière suivante :

$$\begin{aligned}
 \|\cdot\|_2 : \mathbb{R}^n &\rightarrow \mathbb{R} \\
 x = (x_1, x_2, \dots, x_n) &\mapsto \|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2},
 \end{aligned}$$

c'est-à-dire qui vérifie :

$$\|x\|_2^2 = x^T x = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Nous avons à présent effectué tous les rappels nécessaires concernant les matrices. En effet, c'est à partir de celles-ci que nous allons pouvoir détailler des méthodes de résolutions de problèmes mathématiques en lien avec le calcul numérique, comme nous l'avons recherché dans les différents programmes d'enseignement. Ainsi, nous développons ci-après les systèmes d'équations linéaires qui nous serviront dans notre illustration.

2.3 Systèmes d'équations linéaires

Dès lors que les différents rappels concernant le calcul matriciel sont effectués, nous pouvons poursuivre notre analyse par les systèmes d'équations linéaires. Pour ce faire, nous nous inspirons des livres de ADAM A. et LOUSBERG F. [7] et de DODGE Y. [11].

Définition 2.13. Un *système linéaire* est un système de m équations linéaires, c'est-à-dire du premier degré, à n inconnues x_1, x_2, \dots, x_n que l'on écrit sous la forme générale :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

ou sous la forme matricielle :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

que l'on note :

$$Ax = b ,$$

où :

- $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ est la matrice du système ;
- $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ est le vecteur des inconnues ;
- $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ est le vecteur des constantes.

Résoudre le système revient à trouver l'ensemble des solutions de ce système, c'est-à-dire l'ensemble des valeurs des inconnues qui vérifient chacune des équations du système.

Pour résoudre un système d'équations linéaires, il existe différentes méthodes. La plus connue, dans le sens où c'est cette méthode qui est la plus utilisée en secondaire, est la *méthode de substitution*. Elle consiste à isoler une inconnue dans l'une des équations et de la remplacer ensuite dans les autres équations du système. Ce procédé est répété plusieurs fois de sorte à obtenir la valeur de chaque inconnue.

Exemple 2.11. Soit le système linéaire suivant :

$$\begin{cases} x_1 + 2x_2 - x_3 = -3 \\ -x_1 + x_2 + 4x_3 = 0 \\ 5x_1 - 2x_2 - 2x_3 = 6 \end{cases}.$$

On isole x_1 dans la première équation, ce qui nous donne :

$$\begin{cases} x_1 = -2x_2 + x_3 - 3 \\ -x_1 + x_2 + 4x_3 = 0 \\ 5x_1 - 2x_2 - 2x_3 = 6 \end{cases}$$

et, en remplaçant l'expression de x_1 obtenue dans les deux autres équations, on a :

$$\begin{aligned} \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ -(-2x_2 + x_3 - 3) + x_2 + 4x_3 = 0 \\ 5(-2x_2 + x_3 - 3) - 2x_2 - 2x_3 = 6 \end{cases} &\Leftrightarrow \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ (2x_2 - x_3 + 3) + x_2 + 4x_3 = 0 \\ (-10x_2 + 5x_3 - 15) - 2x_2 - 2x_3 = 6 \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ 3x_2 + 3x_3 = -3 \\ -12x_2 + 3x_3 = 21 \end{cases}. \end{aligned}$$

On isole x_2 dans la deuxième équation, ce qui nous donne :

$$\begin{cases} x_1 = -2x_2 + x_3 - 3 \\ 3x_2 = -3x_3 - 3 \\ -12x_2 + 3x_3 = 21 \end{cases} \Leftrightarrow \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ x_2 = -x_3 - 1 \\ -12x_2 + 3x_3 = 21 \end{cases}$$

et, en remplaçant l'expression de x_2 obtenue dans la troisième équation, on obtient :

$$\begin{aligned} \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ x_2 = -x_3 - 1 \\ -12(-x_3 - 1) + 3x_3 = 21 \end{cases} &\Leftrightarrow \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ x_2 = -x_3 - 1 \\ 12x_3 + 3x_3 = 9 \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ x_2 = -x_3 - 1 \\ 15x_3 = 9 \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = -2x_2 + x_3 - 3 \\ x_2 = -x_3 - 1 \\ x_3 = \frac{3}{5} \end{cases} . \end{aligned}$$

On peut à présent remplacer la valeur de x_3 dans la deuxième équation pour obtenir la valeur de x_2 :

$$x_2 = -x_3 - 1 = -\frac{3}{5} - 1 = -\frac{8}{5}$$

et il nous reste à remplacer les valeurs de x_2 et x_3 dans la première équation pour trouver la valeur de x_1 :

$$x_1 = -2x_2 + x_3 - 3 = -2 \cdot \frac{-8}{5} + \frac{3}{5} - 3 = \frac{4}{5} .$$

Remarquons que, à la dernière étape du procédé ci-dessus, nous avons obtenu un système équivalent au système de départ qui peut s'écrire de la manière suivante :

$$\begin{cases} x_1 + 2x_2 - x_3 = -3 \\ x_2 + x_3 = -1 \\ x_3 = \frac{3}{5} \end{cases}$$

ou, avec la notation matricielle :

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \\ \frac{3}{5} \end{pmatrix} .$$

On peut ainsi constater que, d'un point de vue matriciel, le principe général de la méthode de substitution est de trouver un système $\tilde{A}x = \tilde{b}$ équivalent au système de départ $Ax = b$ et tel que \tilde{A} soit une matrice triangulaire dont les éléments diagonaux sont unitaires.

En pratique, lorsque les systèmes linéaires sont de taille trop importante (beaucoup d'inconnues et/ou beaucoup d'équations), il devient difficile de résoudre de tels problèmes « à la main », c'est pourquoi nous devons passer à une résolution numérique, via des ordinateurs. Nous en reparlerons dans le chapitre 3.

2.4 Dérivées et équations aux dérivées partielles

Nous détaillons dans cette section les notions de dérivées de fonctions et pouvons ainsi étendre les notions d'équations linéaires à un cas particulier plus complexe - les équations aux dérivées partielles - bien que celles-ci ne fassent pas partie du programme. Nous en expliquons donc le principe général puisque celui-ci nous sera utile pour la suite de ce travail. Les définitions de cette section proviennent du livre de ANNOYE M. et VAN EERDENBRUGGHE A. [9] ainsi que du cours de LEMAÎTRE A. [3].

Définition 2.14. Soient une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ et un réel a du domaine de définition de f . La fonction f est *continue en a* si et seulement si :

$$\lim_{x \rightarrow a} f(x) = f(a) .$$

Définition 2.15. Soient une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ et un réel a du domaine de définition de f . La fonction f est *dérivable en a* si et seulement si la limite :

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

existe dans \mathbb{R} . Alors, on note $f'(a)$ la valeur de cette limite.

Le *domaine de dérivabilité* de f est l'ensemble des réels a tels que f est dérivable en a :

$$\text{dom}_d f = \{a \in \mathbb{R} \mid f'(a) \exists\} .$$

La *fonction dérivée* de f est la fonction f' définie par :

$$\begin{aligned} f' : \text{dom}_d f &\rightarrow \mathbb{R} \\ x &\mapsto f'(x) \end{aligned} .$$

Si f est dérivable sur tout son domaine de définition, on note $f \in \mathcal{C}^1$.

Exemple 2.12. Soit la fonction $f(x) = x^2 + 3x$. On a :

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{(x+h)^2 + 3(x+h) - (x^2 + 3x)}{h} \\ &= \frac{x^2 + h^2 + 2hx + 3x + 3h - x^2 - 3x}{h} \\ &= \frac{h(h + 2x + 3)}{h} \\ &= h + 2x + 3 \end{aligned}$$

d'où :

$$f'(x) = \lim_{h \rightarrow 0} (h + 2x + 3) = 2x + 3 .$$

Remarque. Soient une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ et un réel $a \in \text{dom}_f$. Si la fonction f est dérivable en a , alors f est continue en a . En effet,

$$\begin{aligned}
 \lim_{x \rightarrow a} f(x) &= \lim_{x \rightarrow a} (f(x) - f(a) + f(a)) \\
 &= \lim_{x \rightarrow a} (f(x) - f(a)) + \lim_{x \rightarrow a} f(a) \\
 &= \lim_{x \rightarrow a} \left(\frac{f(x) - f(a)}{x - a} (x - a) \right) + \lim_{x \rightarrow a} f(a) \\
 &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \rightarrow a} (x - a) + \lim_{x \rightarrow a} f(a) \\
 &= f'(a) \cdot 0 + f(a) \\
 &= f(a) .
 \end{aligned}$$

Les dérivées partielles d'une fonction ne sont pas étudiées dans le cadre de l'enseignement secondaire. Cependant, les élèves connaissant la notion de dérivée peuvent comprendre celle de dérivée partielle, qui nous sera utile pour la suite de ce travail.

Définition 2.16. Soit une fonction de plusieurs variables :

$$\begin{aligned}
 f : \mathbb{R}^n &\rightarrow \mathbb{R} \\
 (x_1, x_2, \dots, x_n) &\mapsto f(x_1, x_2, \dots, x_n) .
 \end{aligned}$$

Les *dérivées partielles* de la fonction f sont les fonctions notées :

$$\frac{\partial f}{\partial x_i} \quad (i = 1, \dots, n)$$

qui correspondent aux dérivées de la fonction calculées en une variable et en considérant les autres variables comme des constantes.

On définit alors le *gradient* de la fonction f de la manière suivante :

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} .$$

Exemple 2.13. Soit la fonction $f(x, y) = x^2 + 3xy$. On a :

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x + 3y \\ 3x \end{pmatrix} .$$

De la même manière, on peut définir les dérivées partielles secondes d'une fonction comme ci-dessous.

Définition 2.17. Soit une fonction de plusieurs variables

$$\begin{aligned} f : \mathbb{R}^n &\rightarrow \mathbb{R} \\ (x_1, x_2, \dots, x_n) &\mapsto f(x_1, x_2, \dots, x_n) \end{aligned}$$

Si f est deux fois dérivable (on note $f \in \mathcal{C}^2$), c'est-à-dire si f est dérivable et si la dérivée de f est elle-même dérivable, alors les *dérivées partielles secondes* de la fonction f sont les fonctions notées :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) \quad (i, j = 1, \dots, n)$$

qui correspondent aux calculs de deux dérivées partielles premières successives de la fonction.

On définit alors la *matrice hessienne* de la fonction f de la manière suivante :

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Exemple 2.14. Soit la fonction $f(x, y) = xy^4 + 3x^2y$. On a :

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} (y^4 + 6xy) & \frac{\partial}{\partial x} (4xy^3 + 3x^2) \\ \frac{\partial}{\partial y} (y^4 + 6xy) & \frac{\partial}{\partial y} (4xy^3 + 3x^2) \end{pmatrix} = \begin{pmatrix} 6y & 4y^3 + 6x \\ 4y^3 + 6x & 12xy^2 \end{pmatrix}.$$

La notion de dérivée d'une fonction peut être considérée « dans l'autre sens », c'est-à-dire que nous pouvons partir de la dérivée d'une fonction pour retrouver la fonction elle-même.

Définition 2.18. Soient une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ et un ensemble $I \subset \text{dom}_f$. La fonction $F : I \rightarrow \mathbb{R}$ est une *primitive* de f sur I si et seulement si F est dérivable sur I et, $\forall x \in I$, $F'(x) = f(x)$.

On appelle *intégrale (indéfinie)* de la fonction f l'ensemble des primitives de f :

$$\int f(x) dx = F(x) + k,$$

où k est une constante réelle quelconque.

Nous pouvons à présent définir la notion d'équation aux dérivées partielles qui est un élément clé dans la résolution de problèmes concrets comme ceux concernant les prévisions météorologiques.

Définition 2.19. Une *équation aux dérivées partielles* (EDP) est une relation d'égalité reliant :

- une fonction inconnue u de n variables ($n \geq 2$) :

$$u(x_1, x_2, \dots, x_n);$$

- une ou plusieurs dérivée(s) partielle(s) de cette fonction u par rapport à ces différentes variables :

$$\frac{\partial^j u}{\partial x_1^{j_1} \partial x_2^{j_2} \dots \partial x_n^{j_n}}, \text{ avec } j_1 + j_2 + \dots + j_n = j;$$

- les variables elles-mêmes ou des fonctions de ces variables.

L'*ordre* d'une EDP correspond à la plus grande valeur de j présente dans celle-ci.

Pour pouvoir trouver une solution d'une EDP, celle-ci s'accompagne de conditions limites. Il s'agit de conditions initiales et/ou de conditions aux frontières du domaine de calcul.

Les EDP les plus répandues sont les EDP linéaires d'ordre 1 et d'ordre 2. La forme générale d'une EDP linéaire d'ordre 1 à deux variables est :

$$a(x, y) \frac{\partial u}{\partial x} + b(x, y) \frac{\partial u}{\partial y} + c(x, y)u = d(x, y),$$

tandis que la forme générale d'une EDP linéaire d'ordre 2 à deux variables est :

$$a(x, y) \frac{\partial^2 u}{\partial x^2} + 2b(x, y) \frac{\partial^2 u}{\partial x \partial y} + c(x, y) \frac{\partial^2 u}{\partial y^2} + d(x, y) \frac{\partial u}{\partial x} + e(x, y) \frac{\partial u}{\partial y} + f(x, y)u = g(x, y),$$

où $a(x, y)$, $b(x, y)$, $c(x, y)$, $d(x, y)$, $e(x, y)$, $f(x, y)$ et $g(x, y)$ sont des fonctions des variables x et y .

Dans la pratique, résoudre des problèmes d'équations aux dérivées partielles n'est pas une tâche aisée. Bien souvent, il n'est pas possible de trouver une solution analytique et, dans ce cas, nous devons modéliser le problème et recourir à une méthode itérative : on part d'une solution initiale et on ajuste l'estimation de la solution recherchée un certain nombre de fois jusqu'à en obtenir une raisonnablement proche de celle-ci. Malheureusement, ce genre de démarche est très souvent irréalisable manuellement et nous avons besoin d'effectuer les calculs sur ordinateur. Dans le chapitre 3, nous verrons comment et dans quelles limites on peut utiliser ces derniers. Dans le chapitre 4, nous verrons que les équations aux dérivées partielles sont à la base des prévisions météorologiques.

2.5 Problèmes d'optimisation sans contrainte

Le cadre général d'un problème d'optimisation sans contrainte² consiste à trouver le minimum ou le maximum d'une fonction. Nous détaillons dans cette section les méthodes de base d'un tel problème. Nous nous inspirons pour cela du cours de STRODIOT J.-J. [6] ainsi que du livre de NOCEDAL J. et WRIGHT S.-J. [14].

2. Un problème d'optimisation avec contrainte consiste à déterminer la plus petite (ou la plus grande) valeur d'une fonction dans un ensemble donné, appelé *ensemble admissible*.

Définition 2.20. Soient une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Un *problème d'optimisation* de f sans contrainte consiste à minimiser (ou maximiser) f sur son domaine :

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{ou} \quad \max_{x \in \mathbb{R}^n} f(x) .$$

De manière générale, on considère uniquement les problèmes de minimisation puisque calculer le maximum d'une fonction revient à calculer l'opposé du minimum de l'opposé de cette fonction.

Un *minimum global* d'un tel problème est un point $x^* \in \mathbb{R}^n$ tel que, $\forall x \in \mathbb{R}^n$, $f(x^*) \leq f(x)$. Un *minimum local* est un point $x^* \in \mathbb{R}^n$ tel qu'il existe un voisinage $N \subset \mathbb{R}^n$ de x^* tel que, pour tout $x \in N$, $f(x^*) \leq f(x)$.

Définition 2.21. Un ensemble C est *convexe* si :

$$\forall x, y \in C, \forall t \in [0, 1], tx + (1 - t)y \in C .$$

Dans ce cas, f est *convexe* si :

$$\forall x, y \in C, \forall t \in [0, 1], f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) .$$

Les théorèmes suivants ne sont pas abordés en secondaire mais sont très utiles pour déterminer la solution d'un problème d'optimisation. Nous nous contentons donc de les citer.

Théorème 2.1. *Le point x^* est un minimum local de f si et seulement si $\nabla^2 f$ est continue dans un voisinage de x^* , $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est semi-définie positive.*

Théorème 2.2. *Si la fonction f est convexe, alors tout minimum local de f est un minimum global de f .*

Remarque. Dans le cas d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, nous avons que x^* est un minimum si $f'(x^*) = 0$ et si, pour x autour de x^* , $f''(x) > 0$. Le théorème 2.1 généralise ce concept.

2.6 Statistiques

Nous allons ici expliquer le principe général de la méthode des moindres carrés appliquée à une série statistique. Nous allons également donner diverses définitions qui nous serviront dans le chapitre 4. Ces définitions sont inspirées du cours de HARDY A. [2] et du livre de ANNOYE M., GILON J.-L., VAN EERDENBRUGGHE A. et WILEMME J. [8].

Commençons par définir les concepts de variance et de covariance.

Définition 2.22. Une *expérience aléatoire* est une expérience dont on peut décrire les résultats possibles mais dont on ne peut pas prédire avec certitude le résultat. On considère qu'on peut reproduire cette expérience dans les mêmes conditions autant de fois que souhaité.

La *catégorie d'épreuve* Ω est l'ensemble des résultats possibles d'une épreuve aléatoire. On appelle *événement* tout sous-ensemble de Ω .

Dans le cadre de la météorologie, comme dans bien d'autres domaines scientifiques, une expérience aléatoire peut être vue comme un phénomène qui évolue au cours du temps et dont on peut obtenir des informations via des observations dans des conditions bien particulières.

Définition 2.23. Une *variable aléatoire* est une fonction de Ω dans \mathbb{R} qui fait correspondre, à chaque résultat d'une expérience aléatoire, une valeur réelle.

Définition 2.24. Soit $x = (x_1, x_2, \dots, x_n)$ une variable aléatoire. La *variance* de cette variable est définie par :

$$Var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 .$$

où $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ est la moyenne des x_i ($i = 1, \dots, n$). L'*écart-type* de la variable x vaut $\sigma(x) = \sqrt{Var(x)}$.

Définition 2.25. Soient $x = (x_1, x_2, \dots, x_n)$ et $y = (y_1, y_2, \dots, y_n)$ deux variables aléatoires. La *covariance* de ces variables est définie par :

$$Cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) .$$

Lorsqu'on dispose de m variables aléatoires, on définit la matrice de covariance de la manière suivante :

Définition 2.26. Soient x, y, z, \dots , des variables aléatoires. La *matrice de covariance* de ces variables est définie par :

$$C_{(m \times m)} = \begin{pmatrix} Var(x) & Cov(x, y) & Cov(x, z) & \cdots \\ Cov(y, x) & Var(y) & Cov(y, z) & \cdots \\ Cov(z, x) & Cov(z, y) & Var(z) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} .$$

Cette matrice est symétrique car, $\forall x$ et $\forall y$,

$$Cov(x, y) = Cov(y, x) .$$

Cette définition nous sera utile dans le chapitre 4. Nous détaillons à présent la méthode des moindres carrés dans le cas d'une régression linéaire. Il s'agit là d'initier le lecteur à l'utilisation des moindres carrés dans la résolution de problèmes.

Soit une série statistique de n variables x_1, x_2, \dots, x_n associées aux valeurs y_1, y_2, \dots, y_n . Si l'on représente ces couples (x_i, y_i) ($i = 1, \dots, n$) dans un repère cartésien, on peut calculer la *droite de régression* linéaire de y par rapport à x en utilisant la méthode des moindres carrés. Le principe de cette méthode est de minimiser la somme des carrés des écarts verticaux des n points donnés aux points de même abscisse sur la droite. Pour ce faire, nous commençons par définir la moyenne d'une série.

La droite de régression aura pour équation $f(x) = ax + b$ et on cherche donc à minimiser la somme :

$$S(a, b) = \sum_{i=1}^n (y_i - (ax_i + b))^2 .$$

Nous devons, dans un premier temps, annuler le gradient de cette fonction de a et b . Pour cela, nous calculons les dérivées partielles de la somme par rapport à a et par rapport à b et les égalons à 0. On obtient le système suivant :

$$\begin{aligned} \nabla S = 0 & \Leftrightarrow \begin{cases} \frac{\partial S}{\partial a} = 0 \\ \frac{\partial S}{\partial b} = 0 \end{cases} \\ & \Leftrightarrow \begin{cases} -2 \sum_{i=1}^n ((y_i - ax_i - b) \cdot x_i) = 0 \\ -2 \sum_{i=1}^n (y_i - ax_i - b) = 0 \end{cases} \\ & \Leftrightarrow \begin{cases} \sum_{i=1}^n ((y_i - ax_i - b) \cdot x_i) = 0 & (1) \\ \sum_{i=1}^n (y_i - ax_i - b) = 0 & (2) \end{cases} \end{aligned}$$

L'équation (2) donne une solution pour b :

$$\begin{aligned} \sum_{i=1}^n (y_i - ax_i - b) = 0 & \Leftrightarrow \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - n \cdot b = 0 \\ & \Leftrightarrow n \cdot b = \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i \\ & \Leftrightarrow b = \frac{1}{n} \sum_{i=1}^n y_i - a \frac{1}{n} \sum_{i=1}^n x_i \\ & \Leftrightarrow b = \bar{y} - a\bar{x} . \end{aligned}$$

Nous pouvons donc remplacer cette valeur dans les équations (1) et (2) et nous obtenons le système suivant :

$$\begin{cases} \sum_{i=1}^n ((y_i - ax_i - \bar{y} + a\bar{x}) \cdot x_i) = 0 \\ \sum_{i=1}^n (y_i - ax_i - \bar{y} + a\bar{x}) = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{i=1}^n (((y_i - \bar{y}) - a(x_i - \bar{x})) \cdot x_i) = 0 & (3) \\ \sum_{i=1}^n ((y_i - \bar{y}) - a(x_i - \bar{x})) = 0 & (4) \end{cases}$$

Finalement, on a :

$$\begin{aligned} (3) - \bar{x} \cdot (4) & \Leftrightarrow \sum_{i=1}^n \left(((y_i - \bar{y}) - a(x_i - \bar{x})) \cdot (x_i - \bar{x}) \right) = 0 \\ & \Leftrightarrow \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}) - a \sum_{i=1}^n (x_i - \bar{x})^2 = 0 \\ & \Leftrightarrow a = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} . \end{aligned}$$

Remarque. On constate que le coefficient a de la droite de régression peut s'écrire comme :

$$a = \frac{Cov(x, y)}{Var(x)}.$$

Il nous reste encore à déterminer si la matrice hessienne de S est semi-définie positive.

$$\begin{aligned}\nabla^2 S &= \begin{pmatrix} \frac{\partial^2 S}{\partial a^2} & \frac{\partial^2 S}{\partial a \partial b} \\ \frac{\partial^2 S}{\partial a \partial b} & \frac{\partial^2 S}{\partial b^2} \end{pmatrix} \\ &= \begin{pmatrix} -2 \cdot \sum_{i=1}^n (-x_i^2) & -2 \cdot \sum_{i=1}^n (-x_i) \\ -2 \cdot \sum_{i=1}^n (-x_i) & -2 \cdot \sum_{i=1}^n (-1) \end{pmatrix} \\ &= \begin{pmatrix} 2 \cdot \sum_{i=1}^n x_i^2 & 2 \cdot \sum_{i=1}^n x_i \\ 2 \cdot \sum_{i=1}^n x_i & 2n \end{pmatrix}\end{aligned}$$

Puisque le déterminant de cette matrice est positif :

$$\begin{aligned}\det(\nabla^2 S) &= 4n \sum_{i=1}^n x_i^2 - 4 \left(\sum_{i=1}^n x_i \right)^2 \\ &= 4n^2 \left(\frac{\sum_{i=1}^n x_i^2}{n} - \frac{\left(\sum_{i=1}^n x_i \right)^2}{n^2} \right) \\ &= 4n^2 \left(\frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 \right) \\ &= 4n^2 Var^2(x) > 0\end{aligned}$$

et que les mineurs principaux sont positifs :

$$2 \sum_{i=1}^n x_i^2 > 0 \text{ et } 2n > 0,$$

nous pouvons appliquer le critère de Sylvester (voir proposition 2.1) et nous avons bien que la matrice hessienne $\nabla^2 S$ est définie positive. Ainsi, les valeurs de a et b obtenues minimisent l'écart entre la droite d'équation $f(x) = ax + b$.

Ce développement de régression linéaire nous montre une application de la méthode des moindres carrés. En fait, on considère les carrés des écarts pour obtenir des valeurs qui soient positives d'une part et, d'autre part, qui accordent une plus grande importance aux valeurs plus éloignées de la moyenne. Nous verrons comment cette méthode peut être appliquée à d'autres types de problèmes mathématiques.

Nous avons à présent passé en revue les matières abordées en secondaire et en lien avec le calcul numérique. Nous avons également étendu notre étude à des notions qui sont d'une importance capitale dans les prévisions météorologiques, que nous aborderons plus loin. Avant cela, nous détaillons les différentes notions concernant le calcul numérique, puisqu'il s'agit de la clé de notre travail.

Chapitre 3

Le calcul numérique

Lorsque les résolutions de problèmes mathématiques, comme les systèmes linéaires ou autres, deviennent trop complexes pour l'être humain, il est utile, voire indispensable, d'avoir recours au calcul sur ordinateur. Celui-ci sera capable de résoudre des problèmes de grande taille dans un laps de temps relativement court. Cependant, il faut savoir l'utiliser, connaître ses capacités et ses limites afin de ne pas commettre d'erreurs. Ces points font l'objet de ce chapitre.

Les informations reprises dans ce chapitre sont inspirées des livres de QUARTERONI A., SACCO R. et SALERI F. [15] et ASCHER U.M., GRIEF C. [10], ainsi que des cours dispensés par FÜZFA A. [1] et par SARTENAER A. [4] et [5].

3.1 Arithmétique en virgule flottante

Chaque ordinateur a sa propre identité, composée d'une base (notée β), d'une précision (notée t), ainsi que d'un intervalle d'exposants, que nous notons $[L, U]$ où, typiquement, $L < 0$ et $U > 0$. Ces différents éléments permettent de définir l'ensemble des nombres que l'ordinateur pourra représenter, appelés *nombres en virgule flottante* :

$$F = \{0\} \cup \{f = \pm .d_1 \dots d_t \times \beta^e; 0 \leq d_i < \beta (d_1 \neq 0); L \leq e \leq U\} ,$$

où $.d_1 \dots d_t$ est une notation abrégée de la *mantisse* $\sum_{i=1}^t d_i \cdot \beta^{-i}$.

Prenons, par exemple, le nombre $x = 4.5$ et tentons de le représenter en code binaire, c'est-à-dire de base 2. Nous avons :

$$\begin{aligned} 4.5 &= 4 + 0.5 \\ &= 2^2 + 2^{-1} \\ &= (2^{-1} + 2^{-4}) \cdot 2^3 , \end{aligned}$$

et x s'écrit donc en représentation binaire de la manière suivante : $x = + .1001 \times 2^3$.

Supposons que l'ordinateur dispose de N cases mémoires, appelées *bits*, pour chacun des nombres stockés. Un nombre en virgule flottante est alors réparti de la manière suivante (voir figure 3.1) :

- une case est consacrée au signe de ce nombre (par convention, cette case sera réservée à un représentant de la forme binaire : 0 si le nombre est positif et 1 s'il est négatif) ;

- la mantisse, ou les paramètres d_i , $i \in \{1, \dots, t\}$, est stockée dans les t cases suivantes ;
- la valeur de l'exposant e est contenue dans les $N - t - 1$ cases restantes.

En revanche, le stockage du nombre 0 se fait d'une manière différente, que nous n'expliquerons pas dans le cadre de ce mémoire.

\pm	d_1, d_2, \dots, d_t	e
-------	------------------------	-----

FIGURE 3.1 – Représentation du stockage des nombres en virgule flottante.

Dans la plupart des ordinateurs, la base β vaut 2. Il existe typiquement deux types de représentation d'un nombre en virgule flottante sur un ordinateur : la représentation en *simple précision* ($N = 32$ bits) et celle en *double précision* ($N = 64$ bits). Dans le cas de la simple précision, $t = 23$ et $e \in [-32, 32]$. Dès lors, 23 bits sont consacrés à la mantisse et il reste donc 8 bits pour l'exposant. Dans le cas de la double précision, $t = 52$ et $e \in [-64, 64]$. Ainsi, 52 bits sont pour la mantisse et 11 bits pour l'exposant.

Avec cette représentation des nombres en virgule flottante, nous pouvons estimer l'erreur d'arrondi d'un nombre encodé en machine en calculant des bornes de celle-ci. La proposition 3.1 permet de déterminer les valeurs maximale et minimale que peut prendre un nombre en virgule flottante, en fonction des valeurs de la base, de la précision et de l'intervalle d'exposants.

Proposition 3.1. *Soit $f \in F$. Si $f \neq 0$, alors $m = \beta^{L-1} \leq |f| \leq (1 - \beta^{-t}) \cdot \beta^U = M$.*

Démonstration.

Le plus petit nombre représentable en machine, en valeur absolue, est donné par :

$$\begin{aligned}
 m &= .100 \dots 0 \times \beta^L \\
 &= (1 \cdot \beta^{-1} + 0 \cdot \beta^{-2} + 0 \cdot \beta^{-3} + \dots + 0 \cdot \beta^{-t}) \times \beta^L \\
 &= \beta^{-1} \times \beta^L \\
 &= \beta^{L-1} .
 \end{aligned}$$

Le plus grand nombre représentable en machine, en valeur absolue, est donné par :

$$\begin{aligned}
 M &= .(\beta - 1)(\beta - 1) \dots (\beta - 1) \times \beta^U \\
 &= ((\beta - 1) \cdot \beta^{-1} + (\beta - 1) \cdot \beta^{-2} + \dots + (\beta - 1) \cdot \beta^{-t}) \times \beta^U \\
 &= \left(\sum_{i=1}^t (\beta - 1) \cdot \beta^{-i} \right) \times \beta^U \\
 &= (1 - \beta^{-t}) \times \beta^U .
 \end{aligned}$$

En effet, nous pouvons affirmer la dernière égalité grâce au développement suivant :

$$\left(\sum_{i=1}^t (\beta - 1) \cdot \beta^{-i} \right) + \beta^{-t} = (\beta - 1) \cdot \beta^{-1} + (\beta - 1) \cdot \beta^{-2} + \dots + (\beta - 1) \cdot \beta^{-t+1} + \underbrace{(\beta - 1) \cdot \beta^{-t} + \beta^{-t}}_{=1}$$

$$\begin{aligned}
&= (\beta - 1) \cdot \beta^{-1} + (\beta - 1) \cdot \beta^{-2} + \dots + (\beta - 1) \cdot \beta^{-t+1} + ((\beta - 1) + 1) \cdot \beta^{-t} \\
&= (\beta - 1) \cdot \beta^{-1} + (\beta - 1) \cdot \beta^{-2} + \dots + \underbrace{(\beta - 1) \cdot \beta^{-t+1} + \beta^{-t+1}}_{\beta^{-t+1}} \\
&= (\beta - 1) \cdot \beta^{-1} + (\beta - 1) \cdot \beta^{-2} + \dots + ((\beta - 1) + 1) \cdot \beta^{-t+1} \\
&= (\beta - 1) \cdot \beta^{-1} + (\beta - 1) \cdot \beta^{-2} + \dots + \beta^{-t+2} \\
&\vdots \\
&= (\beta - 1) \cdot \beta^{-1} + \underbrace{(\beta - 1) \cdot \beta^{-2} + \beta^{-2}}_{\beta^{-2}} \\
&= (\beta - 1) \cdot \beta^{-1} + ((\beta - 1) + 1) \cdot \beta^{-2} \\
&= \underbrace{(\beta - 1) \cdot \beta^{-1} + \beta^{-1}}_{\beta^{-1}} \\
&= ((\beta - 1) + 1) \cdot \beta^{-1} \\
&= \beta^0 \\
&= 1,
\end{aligned}$$

c'est-à-dire $\sum_{i=1}^t (\beta - 1) \cdot \beta^{-i} = 1 - \beta^{-t}$.

■

À titre d'exemple, prenons $\beta = 2$, $t = 3$ et $[L, U] = [0, 2]$. Nous obtenons ainsi

$$F = \pm \left\{ \begin{array}{l} 0 \\ \cdot 100 \times \left\{ \begin{array}{l} 2^0 \\ 2^1 \\ 2^2 \end{array} \right. \\ \cdot 110 \times \left\{ \begin{array}{l} 2^0 \\ 2^1 \\ 2^2 \end{array} \right. \\ \cdot 101 \times \left\{ \begin{array}{l} 2^0 \\ 2^1 \\ 2^2 \end{array} \right. \\ \cdot 111 \times \left\{ \begin{array}{l} 2^0 \\ 2^1 \\ 2^2 \end{array} \right. \end{array} \right. = \pm \left\{ \begin{array}{l} 0 \\ (1 \cdot 2^{-1}) \cdot 2^0 \\ (1 \cdot 2^{-1}) \cdot 2^1 \\ (1 \cdot 2^{-1}) \cdot 2^2 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-2}) \cdot 2^0 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-2}) \cdot 2^1 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-2}) \cdot 2^2 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-3}) \cdot 2^0 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-3}) \cdot 2^1 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-3}) \cdot 2^2 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-2} + 1 \cdot 2^{-3}) \cdot 2^0 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-2} + 1 \cdot 2^{-3}) \cdot 2^1 \\ (1 \cdot 2^{-1} + 1 \cdot 2^{-2} + 1 \cdot 2^{-3}) \cdot 2^2 \end{array} \right. = \pm \left\{ \begin{array}{l} 0 \\ 0.5 \\ 1 \\ 2 \\ 0.75 \\ 1.5 \\ 3 \\ 0.625 \\ 1.25 \\ 2.5 \\ 0.875 \\ 1.75 \\ 3.5 \end{array} \right.$$

et nous pouvons vérifier les valeurs des plus grand et plus petit nombres représentables en machine (en valeur absolue) :

$$\begin{aligned}
m &= \beta^{L-1} = 2^{0-1} = 2^{-1} = 0.5, \\
M &= (1 - \beta^{-t}) \cdot \beta^U = (1 - 2^{-3}) \cdot 2^2 = 3.5.
\end{aligned}$$

La représentation graphique de F pour cet exemple se trouve à la Figure 3.2. Nous constatons que la répartition des différents nombres en virgule flottante se fait de manière irrégulière.

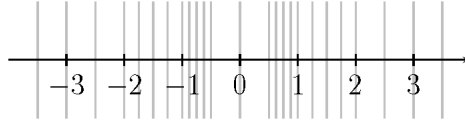


FIGURE 3.2 – Représentation de l'ensemble F avec $\beta = 2$, $t = 3$ et $[L, U] = [0, 2]$.

Définissons à présent l'ensemble :

$$G = \{0\} \cup \{x \in \mathbb{R} \text{ tel que } m \leq |x| \leq M\} ,$$

et l'opérateur :

$$\begin{aligned} fl : G &\longrightarrow F \\ x &\longmapsto fl(x) = c \end{aligned} ,$$

où c est le plus proche élément de x dans F (en arrondissant au plus loin de 0). La proposition 3.2 détermine des bornes d'erreur lors de l'approximation d'un nombre réel par un nombre en virgule flottante.

Proposition 3.2. $\forall x \in G$, $fl(x) = x \cdot (1 + \epsilon)$, avec $|\epsilon| \leq u$, où $u = \frac{1}{2} \cdot \beta^{1-t}$ est la précision machine.

Ceci signifie que chaque nombre réel nul ou compris entre m et M en valeur absolue sera représenté en machine par un nombre éloigné de celui-ci d'une valeur maximale, en valeur absolue, de u . Cette valeur u dépend de la base β et de la précision t de l'ordinateur.

Démonstration.

- si $x = 0$, alors $fl(x) = 0$ et le résultat est immédiat.
- si $x = \pm .d_1 \dots d_t d_{t+1} \times \beta^{e(x)} \neq 0$, alors nous pouvons écrire :

$$fl(x) = x \cdot (1 + \epsilon) \Leftrightarrow \epsilon = \frac{fl(x) - x}{x} ,$$

et nous avons :

$$|\epsilon| = \frac{|fl(x) - x|}{|x|} .$$

Or, d'une part, x , en valeur absolue, peut être minoré de la manière suivante :

$$\begin{aligned} |x| &\geq .10 \dots 0 \times \beta^e \\ &= \beta^{-1} \times \beta^e \\ &= \frac{\beta^e}{\beta} , \end{aligned}$$

d'où :

$$\frac{1}{|x|} \leq \frac{\beta}{\beta^e} ,$$

et, d'autre part, $fl(x) - x$, en valeur absolue, peut être majoré de la manière suivante :

$$\begin{aligned} |fl(x) - x| &= |(\cdot d_1 \dots d_t \times \beta^e) - (\cdot d_1 \dots d_t d_{t+1} \dots \times \beta^e)| \\ &\leq \frac{1}{2} \times \cdot 0 \dots 01 \times \beta^e \\ &= \frac{1}{2} \beta^{-t} \beta^e . \end{aligned}$$

Finalement, nous obtenons :

$$|\epsilon| = \frac{|fl(x) - x|}{|x|} \leq \frac{1}{2} \beta^{-t} \beta^e \frac{\beta}{\beta^e} = \frac{1}{2} \beta^{1-t} .$$

■

Reprenons l'exemple ci-dessus, avec $\beta = 2$, $t = 3$ et $[L, U] = [0, 2]$. Nous obtenons :

$$u = \frac{1}{2} \cdot 2^{1-3} = 2^{-3} = \frac{1}{8} = 0.125 .$$

D'après la proposition 3.2, chaque nombre réel nul ou compris entre 0.5 et 3.5 en valeur absolue sera donc représenté en machine par un nombre éloigné de celui-ci d'au plus 0.125. Si nous prenons, par exemple, le réel $x = 2.75$, nous obtenons :

$$fl(x) = fl(2.75) = 3 ,$$

puisque 3 est la valeur représentable en machine la plus proche de 2.75, et l'erreur relative entre ces nombres vaut :

$$\frac{|3 - 2.75|}{|2.75|} \simeq 0.0909 \leq 0.125 .$$

3.2 Opérations et erreurs d'arrondi

Soient deux nombres en virgule flottante : a et $b \in F$. Notons \odot l'une des opérations arithmétiques élémentaires (addition, soustraction, multiplication ou division) sur ces nombres. Ce type d'opération est appelé *flop* (*floating point operation*).

Si $a \odot b \notin G$, alors on dit qu'il y a *overflow* si $|a \odot b| > M$ ou qu'il y a *underflow* si $|a \odot b| < m$. Par contre, si $a \odot b \in G$, alors nous pouvons supposer que la valeur calculée de $a \odot b$ sera donnée par :

$$fl(a \odot b) = (a \odot b)(1 + \epsilon) ,$$

avec $|\epsilon| \leq u$. À chaque opération effectuée s'ajoute donc une erreur d'arrondi. Une longue suite d'opérations risque ainsi de provoquer de grandes erreurs dans le résultat final. Néanmoins, il est généralement possible d'éviter cela en reformulant le problème ou en associant les opérations dans un ordre pertinent.

3.3 Algorithmes

Définition 3.1. Un *algorithme* est un ensemble d'instructions spécifiant une suite d'opérations élémentaires qui, lorsqu'elles sont exécutées, donnent la solution d'un problème en un nombre fini d'étapes.

La construction d'un algorithme se fait en quatre étapes :

1. formulation du problème ;
2. recherche de la solution mathématique du problème ;
3. implémentation de la résolution du problème sous forme d'algorithme ;
4. vérification et validation de l'algorithme.

Afin de pouvoir comparer et classer différents algorithmes servant à la résolution d'un même problème, nous nous basons sur différents critères comme la *complexité* (nombre d'opérations nécessaires à l'exécution), la *quantité de stockage* (place occupée en mémoire) et les *bornes d'erreur* (estimation de l'erreur dans les résultats obtenus).

Typiquement, on définit la complexité d'un algorithme de manière qualitative, et non quantitative. On dit alors, par exemple, qu'un algorithme est « de l'ordre de n » (que l'on note $\mathcal{O}(n)$) ou « linéaire ». Le tableau ci-dessous reprend les principales complexités par ordre croissant.

Ordre de grandeur	Type de complexité
$\mathcal{O}(1)$	Constante
$\mathcal{O}(\log_2(n))$	Logarithmique
$\mathcal{O}(n)$	Linéaire
$\mathcal{O}(n \log_2(n))$	Quasi-linéaire
$\mathcal{O}(n^2)$	Quadratique
$\mathcal{O}(n^3)$	Cubique
$\mathcal{O}(n^k)$, $k > 2$	Polynomiale
$\mathcal{O}(n^{\log_2(n)})$	Quasi-polynomiale
$\mathcal{O}(2^n)$	Exponentielle
$\mathcal{O}(n!)$	Factorielle

TABLE 3.1 – Complexités d'un algorithme.

Définition 3.2. Un algorithme est dit *stable* s'il n'amplifie pas de trop les erreurs d'arrondi. Un algorithme est dit *instable* s'il amplifie beaucoup les erreurs d'arrondi.

Prenons l'exemple du calcul du produit matriciel. Rappelons que le produit de deux matrices $A_{(m \times p)} = (a_{ij})$ et $B_{(p \times n)} = (b_{ij})$ est la matrice $C_{(m \times n)} = (c_{ij})$ définie par :

$$c_{ij} = \sum_{k=1}^p a_{ik} \cdot b_{kj} ,$$

où $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$ (chapitre 2, définition 2.4). L'idée de l'algorithme est la suivante :

```

| faire pour i=1 jusqu'à m
| | faire pour j=1 jusqu'à n
| | C(i,j) = 0
| | | faire pour k=1 jusqu'à p
| | | C(i,j) = C(i,j) + A(i,k)*B(k,j)
| | | fin de la boucle sur k
| | fin de la boucle sur j
| fin de la boucle sur i

```

Nous avons ainsi une complexité cubique puisqu'il y a $2pnm$ flops. La quantité de stockage est, quant à elle, quadratique car $mn + mp + pn$ données sont mémorisées. Nous pouvons calculer les bornes d'erreur de la manière suivante :

Posons $s_{qij} = fl\left(\sum_{k=1}^q a_{ik} \cdot b_{kj}\right) \forall i \in \{1, \dots, m\}$ et $\forall j \in \{1, \dots, n\}$. Nous avons ainsi :

$$\begin{aligned}
\text{pour } q = 1, \quad s_{1ij} &= a_{i1}b_{1j}(1 + \delta_{1ij}) , \text{ où } |\delta_{1ij}| \leq u \\
\text{pour } q = 2, \dots, n, \quad s_{qij} &= fl(s_{q-1ij} + fl(a_{iq}b_{qj})) \\
&= (s_{q-1ij} + a_{iq}b_{qj}(1 + \delta_{qij}))(1 + \epsilon_{qij}) , \text{ où } |\delta_{qij}| \leq u \text{ et } |\epsilon_{qij}| \leq u
\end{aligned}$$

et :

$$fl((A \cdot B)_{ij}) = s_{p_{ij}} = \sum_{k=1}^p a_{ik}b_{kj}(1 + \gamma_{kij})$$

où $(1 + \gamma_{kij}) = (1 + \delta_{kij}) \prod_{l=k}^p (1 + \epsilon_{lij})$, avec $\epsilon_{1ij} = 0 \forall i \in \{1, \dots, m\}$ et $\forall j \in \{1, \dots, n\}$.

Dès lors,

$$\begin{aligned}
|fl((A \cdot B)_{ij}) - (A \cdot B)_{ij}| &= \left| \sum_{k=1}^p a_{ik}b_{kj}(1 + \gamma_{kij}) - \sum_{k=1}^p a_{ik}b_{kj} \right| \\
&= \left| \sum_{k=1}^p a_{ik}b_{kj}\gamma_{kij} \right| \\
&\leq \sum_{k=1}^p |a_{ik}||b_{kj}||\gamma_{kij}| .
\end{aligned}$$

Or, un résultat du calcul numérique nous permet d'obtenir :

$$|\gamma_{kij}| \leq nu + \mathcal{O}(u^2) ,$$

et donc :

$$\begin{aligned}
|fl((A \cdot B)_{ij}) - (A \cdot B)_{ij}| &\leq \sum_{k=1}^p |a_{ik}||b_{kj}|(nu + \mathcal{O}(u^2)) \\
&= nu|A||B| + \mathcal{O}(u^2) .
\end{aligned}$$

3.4 Illustrations

3.4.1 Résolution de systèmes d'équations

Éliminations de Gauss

Nous prenons ici l'exemple d'un problème de résolution de système d'équations. L'idée est celle développée dans le chapitre 2, à la section 2.3. Pour rappel, le système s'écrit, sous forme matricielle, comme :

$$Ax = b ,$$

où A est une matrice carrée d'ordre n , $x \in \mathbb{R}^n$ et $b \in \mathbb{R}^n$.

Nous avons vu dans le chapitre 2 un exemple très simple dans lequel $n = 3$. Nous avons vu également que la méthode, dite de substitution, consiste, d'un point de vue matriciel, à reformuler le système par un autre système équivalent - c'est-à-dire qui admet la même solution - et dont la matrice est triangulaire. Il existe un algorithme qui généralise cette méthode afin de pouvoir l'appliquer à des dimensions bien plus élevées.

On commence par stocker la matrice du système ainsi que le vecteur des constantes dans une seule matrice $(A|b)_{(n \times (n+1))}$:

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array} \right) .$$

Ensuite, on transforme cette matrice en annulant les termes sous la diagonale; colonne par colonne. On appelle cela les *éliminations de Gauss*.

La première étape, à l'itération k , consiste à diviser chaque élément situé sous l'élément a_{kk} , appelé *pivot*, par celui-ci pour obtenir $l(i, k) = \frac{a_{ik}}{a_{kk}}$. Notons que si le pivot est nul, nous ne pouvons pas appliquer l'algorithme et d'autres méthodes doivent être développées. Par ailleurs, s'il est d'une valeur très petite, l'algorithme peut être instable. Ensuite, pour annuler l'élément a_{ik} (où $i = k + 1, \dots, n$), élément sous la diagonale de la colonne k , il suffit de lui retrancher la valeur :

$$a_{ik} = \frac{a_{ik}}{a_{kk}} \cdot a_{kk} = l(i, k) \cdot a_{kk} ,$$

et, pour obtenir un système équivalent, nous appliquons la même transformation à tous les éléments de la ligne :

$$a_{ij} = a_{ij} - \frac{a_{ik}}{a_{kk}} \cdot a_{kj} = a_{ij} - l(i, k) \cdot a_{kj} ,$$

y compris l'élément $b(i)$:

$$b_i = b_i - l(i, k) \cdot b_k .$$

On continue ce processus pour chacune des colonnes de la matrice. On obtient donc l'algorithme suivant :

```

| faire pour k=1 jusque n-1
| | faire pour i=k+1 jusque n
| | l(i,k) = A(i,k)/A(k,k)
| | | faire pour j=k+1 jusque n
| | | A(i,j) = A(i,j)-l(i,k)*A(k,j)
| | | fin de la boucle sur j
| | b(i) = b(i)-l(i,k)*b(k)
| | fin de la boucle sur i
| fin de la boucle sur k

```

Cet algorithme nécessite un nombre d'opérations qui dépend de n :

$$\#flops = \sum_{k=1}^{n-1} ((n-k) + 2(n-k)(n-k+1)) \simeq \frac{2}{3}n^3 + \mathcal{O}(n^2) .$$

Pour illustrer celui-ci, nous revenons à l'exemple 2.11 du chapitre 2. Soit donc :

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 2 & -1 & -3 \\ -1 & 1 & 4 & 0 \\ 5 & -2 & -2 & 6 \end{array} \right) .$$

On a :

— $k = 1$

- $i = 2 \rightarrow l(2,1) = \frac{A(2,1)}{A(1,1)} = \frac{-1}{1} = -1$

- $j = 2 \rightarrow A(2,2) = A(2,2) - l(2,1) \cdot A(1,2) = 1 - (-1) \cdot 2 = 3$

- $j = 3 \rightarrow A(2,3) = A(2,3) - l(2,1) \cdot A(1,3) = 4 - (-1) \cdot (-1) = 3$

- $b(2) = b(2) - l(2,1) \cdot b(1) = 0 - (-1) \cdot (-3) = -3$

- $i = 3 \rightarrow l(3,1) = \frac{A(3,1)}{A(1,1)} = \frac{5}{1} = 5$

- $j = 2 \rightarrow A(3,2) = A(3,2) - l(3,1) \cdot A(1,2) = -2 - 5 \cdot 2 = -12$

- $j = 3 \rightarrow A(3,3) = A(3,3) - l(3,1) \cdot A(1,3) = -2 - 5 \cdot (-1) = 3$

- $b(3) = b(3) - l(3,1) \cdot b(1) = 6 - 5 \cdot (-3) = 21$

— $k = 2$

- $i = 3 \rightarrow l(3,2) = \frac{A(3,2)}{A(2,2)} = \frac{-12}{3} = -4$

- $j = 3 \rightarrow A(3,3) = A(3,3) - l(3,2) \cdot A(2,3) = 3 - (-4) \cdot 3 = 15$

- $b(3) = b(3) - l(3,2) \cdot b(2) = 21 - (-4) \cdot (-3) = 9$

On obtient, de cette manière, le système modifié :

$$(\bar{A}|\bar{b}) = \left(\begin{array}{ccc|c} 1 & 2 & -1 & -3 \\ 0 & 3 & 3 & -3 \\ 0 & 0 & 15 & 9 \end{array} \right) ,$$

qui correspond au système d'équations suivant :

$$\begin{cases} x_1 + 2x_2 - x_3 = -3 \\ 3x_2 + 3x_3 = -3 \\ 15x_3 = 9 \end{cases} \Leftrightarrow \begin{cases} x_1 + 2x_2 - x_3 = -3 \\ x_2 + x_3 = -1 \\ x_3 = \frac{3}{5} \end{cases}$$

et on constate que l'on retrouve exactement le système de l'exemple 2.11.

Une fois les éliminations de Gauss effectuées, il ne reste plus qu'à procéder à un algorithme, dit de *substitution régressive*, qui est le suivant :

```

b(n) = b(n)/A(n,n)
| faire pour i=n-1 jusque 1 par pas de -1
| c=0
| | faire pour j=i+1 jusque n
| | c = c+A(i,j)*b(j)
| | fin de la boucle sur j
| b(i) = [b(i)-c]/A(i,i) fin de la boucle sur i

```

On constate que les éléments du vecteur b sont remplacés par d'autres valeurs, qui sont alors les solutions du système. De cette manière, on gagne en stockage de données. Le nombre d'opérations nécessaires pour la résolution de l'algorithme de substitution régressive est :

$$\#flops = 1 + \sum_{k=1}^{n-1} ((n-k) + (n-k) + 1) \simeq n^2 .$$

Reprenons notre système $(\bar{A}|\bar{b})$ et appliquons l'algorithme ci-dessus. On a ainsi :

$$\begin{aligned}
& \text{--- } b(3) = \frac{b(3)}{A(3,3)} = \frac{9}{15} = \frac{3}{5} \\
& \text{--- } i = 2 : \\
& \quad \bullet c = 0 \\
& \quad \bullet j = 3 : c = 0 + A(2,3) \cdot b(3) = 3 \cdot \frac{3}{5} = \frac{9}{5} \\
& \rightarrow b(2) = \frac{b(2) - c}{A(2,2)} = \frac{-3 - \frac{9}{5}}{3} = \frac{-8}{5} \\
& \text{--- } i = 1 : \\
& \quad \bullet c = 0 \\
& \quad \bullet j = 2 : c = 0 + A(1,2) \cdot b(2) = 2 \cdot \frac{-8}{5} = \frac{-16}{5} \\
& \quad \bullet j = 3 : c = \frac{-16}{5} + A(1,3) \cdot b(3) = \frac{-16}{5} + (-1) \cdot \frac{3}{5} = \frac{-19}{5} \\
& \rightarrow b(1) = \frac{b(1) - c}{A(1,1)} = \frac{-3 - \frac{-19}{5}}{1} = \frac{4}{5}
\end{aligned}$$

La solution est bien celle obtenue dans le chapitre 2.

La méthode d'éliminations de Gauss est très intéressante car elle permet de générer la solution d'un système d'équations en un nombre d'opérations plutôt intéressant. Cependant, l'inconvénient de celle-ci est que, si le vecteur des constantes b est modifié, tout

le processus doit être répété depuis le début. Pour remédier à ce problème, il existe une autre méthode de résolution de systèmes d'équations qui modifie la matrice du système sans modifier le vecteur des constantes. Il s'agit de la *factorisation LU*.

Factorisation LU

Commençons par définir une matrice carrée d'ordre n , appelée *transformation de Gauss*, par :

$$M_k = I_n - \tau \cdot e_k^T$$

où I_n est la matrice identité d'ordre n (voir définition 2.4) et où τ , appelé *vecteur de Gauss*, et e_k sont définis par :

$$\tau = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \tau_{k+1} \\ \vdots \\ \tau_n \end{pmatrix} \leftarrow k^{\text{ième}} \text{ ligne}, \text{ avec } \tau_i = \frac{x_i}{x_k} \ (\forall i = k+1, \dots, n), \text{ et } e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k^{\text{ième}} \text{ ligne}.$$

On a donc :

$$\begin{aligned} M_k &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \tau_{k+1} \\ \vdots \\ \tau_n \end{pmatrix} \cdot \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & \dots & 0 \\ 0 & 1 & & \vdots & & & \vdots \\ \vdots & & \ddots & 0 & & & \\ & & & 1 & & & \\ & & & -\tau_{k+1} & 1 & & \\ & & & \vdots & & \ddots & 0 \\ 0 & \dots & & -\tau_n & \dots & 0 & 1 \end{pmatrix} \leftarrow k+1^{\text{ième}} \text{ ligne} \\ &\quad \uparrow \\ &\quad k^{\text{ième}} \text{ ligne} \end{aligned}$$

Cette matrice permet de factoriser la matrice du système A en un produit de matrices triangulaires :

$$A = L \cdot U$$

où $L = M_1^{-1} \cdot M_2^{-1} \cdot \dots \cdot M_{n-1}^{-1}$ est triangulaire inférieure avec les éléments diagonaux unitaires et où U , qui est en général égal à $M_{n-1} \cdot \dots \cdot M_2 \cdot M_1 \cdot A$, est triangulaire supérieure. Ainsi, résoudre le système $Ax = b$ revient à résoudre successivement :

$$\begin{cases} Ly = b \\ Ux = y \end{cases}.$$

L'algorithme de factorisation LU de A est donné par :

```

k=1
| faire tant que (A(k,k) != 0) & (k < n )
| | faire pour i=k+1 jusque n
| | | A(i,k) = A(i,k) / A(k,k) % stockage des multiplicateurs  $\tau_i$ 
| | | faire pour j=k+1 jusque n
| | | | A(i,j) = A(i,j) - A(i,k)*A(k,j) % mise à jour de la sous-matrice
| | | fin de la boucle sur j
| | fin de la boucle sur i
| k → k+1

```

On constate que les éléments de la matrice A sont remplacés par d'autres valeurs. En fait, il s'agit de stocker les matrices L et U dans A afin d'éviter un stockage de données supplémentaire. Le nombre d'opérations nécessaires pour la résolution de cet algorithme est :

$$\#flops = \sum_{k=1}^{n-1} \left((n - (k + 1) + 1) + 2(n - (k + 1) + 1)^2 \right) \simeq \frac{2}{3}n^3.$$

Pour illustrer cet algorithme, reprenons l'exemple 2.11 du chapitre 2. Nous avons le système :

$$\begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 4 \\ 5 & -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 6 \end{pmatrix}.$$

Appliquons l'algorithme à ce système. On a :

- $k = 1 < 3 : A(1,1) = 1 \neq 0$
 - $i = 2 \rightarrow A(2,1) = \frac{A(2,1)}{A(1,1)} = \frac{-1}{1} = -1$
 - $j = 2 \rightarrow A(2,2) = A(2,2) - A(2,1) \cdot A(1,2) = 1 - (-1) \cdot 2 = 3$
 - $j = 3 \rightarrow A(2,3) = A(2,3) - A(2,1) \cdot A(1,3) = 4 - (-1) \cdot (-1) = 3$
 - $i = 3 \rightarrow A(3,1) = \frac{A(3,1)}{A(1,1)} = \frac{5}{1} = 5$
 - $j = 2 \rightarrow A(3,2) = A(3,2) - A(3,1) \cdot A(1,2) = -2 - 5 \cdot 2 = -12$
 - $j = 3 \rightarrow A(3,3) = A(3,3) - A(3,1) \cdot A(1,3) = -2 - 5 \cdot (-1) = 3$
- $k = 2 < 3 : A(2,2) = 3 \neq 0$

- $i = 3 \rightarrow A(3, 2) = \frac{A(3, 2)}{A(2, 2)} = \frac{-12}{3} = -4$
- $j = 3 \rightarrow A(3, 3) = A(3, 3) - A(3, 2) \cdot A(2, 3) = 3 - (-4) \cdot 4 = 19$
- $k = 3 \Rightarrow \text{STOP}$.

On obtient ainsi la matrice suivante :

$$\begin{pmatrix} 1 & 2 & -1 \\ -1 & 3 & 3 \\ 5 & -4 & 19 \end{pmatrix}$$

La partie supérieure de cette matrice (y compris la diagonale) représente la partie triangulaire supérieure de la matrice U alors que la partie triangulaire inférieure correspond à la partie triangulaire inférieure (non compris la diagonale qui est, rappelons-le, unitaire) de la matrice L :

$$U = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 19 \end{pmatrix} \text{ et } L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 5 & -4 & 1 \end{pmatrix}$$

Vérifions que le produit de ces matrices donne la matrice A :

$$L \cdot U = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 5 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 4 \\ 0 & 0 & 11 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 4 \\ 5 & -2 & -2 \end{pmatrix}.$$

Il reste donc à résoudre deux systèmes linéaires : un avec une matrice triangulaire inférieure et l'autre avec une matrice triangulaire supérieure. Pour cela, nous avons l'algorithme de substitution régressive, donné au point précédent, ainsi que l'algorithme de substitution progressive suivant :

```

b(1) = b(1)/A(1,1)
| faire pour i=2 jusque n
| c=0
| | faire pour j=1 jusque i-1
| | c = c+A(i,j)*b(j)
| | fin de la boucle sur j
| b(i) = [b(i)-c]/A(i,i)
| fin de la boucle sur i

```

On constate, tout comme pour la substitution régressive, que les éléments du vecteur b sont remplacés par d'autres valeurs. Le nombre d'opérations nécessaires pour la résolution de l'algorithme de substitution progressive est :

$$\#flops = 1 + \sum_{i=2}^{n-1} (1 + 2(i-1)) \simeq n^2.$$

Nous pouvons illustrer cet algorithme en utilisant, à nouveau, l'exemple 2.11 et nous obtenons la même solution qu'avec la méthode d'éliminations de Gauss. Nous laissons cet exercice au lecteur.

Remarque. La méthode de factorisation LU permet d'effectuer des modifications sur la matrice du système A uniquement. Ainsi, il n'y a aucune hypothèse sur le vecteur des constantes b . Cependant, notons qu'il faut être vigilant quant à la structure de la matrice A . En effet, la factorisation LU n'est possible que sous certaines conditions, dont nous ne parlerons pas dans ce mémoire.

Nous avons vu dans ce point comment les résolutions de systèmes linéaires peuvent être réalisées à partir du calcul numérique. Il existe, évidemment, bien d'autres méthodes de résolution mais nous n'en avons présenté que deux qui, selon nous, peuvent être comprises par les élèves du secondaire. Dans le point suivant, nous montrons un autre type de résolution de problème via le calcul numérique. Il s'agit de problèmes d'optimisation.

3.4.2 Optimisation sans contrainte : Méthode de Newton et problèmes aux moindres carrés

La méthode de Newton en optimisation sans contrainte consiste à déterminer la solution x^* d'un problème du type :

$$\min_{x \in \mathbb{R}^n} f(x) ,$$

où f est une fonction de \mathbb{R}^n dans \mathbb{R} . Nous avons abordé ce type de problème dans la section 2.5 du chapitre 2. Nous y avons vu comment en trouver la solution. Cependant, il est bien souvent impossible de trouver une solution analytique à ce type de problème et nous devons alors trouver une approximation raisonnable de cette solution.

Il s'agit donc ici d'une méthode itérative qui détermine, à chaque itération, une nouvelle estimation de x^* :

$$x_{k+1} = x_k + \alpha_k p_k \text{ tel que } f(x_{k+1}) < f(x_k) ,$$

où $p_k \in \mathbb{R}^n$ est appelé *direction de descente* et $\alpha_k \in \mathbb{R}$ est appelé *pas de descente*. Ce vecteur est la solution du problème suivant :

$$\min_{\alpha > 0} f(x_k + \alpha p_k) .$$

L'idée d'une telle méthode est donc de démarrer d'un point quelconque de l'espace \mathbb{R}^n et de « descendre » l'image de ce point à chaque itération, jusqu'à approcher suffisamment la valeur minimale.

La méthode de Newton consiste à prendre comme direction de descente le vecteur :

$$p_k = - \left(\nabla^2 f(x_k) \right)^{-1} \nabla f(x_k) .$$

En effet, ce résultat s'obtient via le théorème suivant, appelé *théorème de Taylor*.

Théorème 3.1. Soit $x = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}^T$ et supposons que $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction au moins trois fois dérivable. Alors, pour $p = \begin{pmatrix} p_1 & p_2 & \dots & p_n \end{pmatrix}^T$, on a :

$$\phi(x + p) = \phi(x) + p^T \nabla \phi(x) + \frac{1}{2} p^T \nabla^2 \phi(x) p + \mathcal{O}(\|p\|^3) .$$

En appliquant ce théorème à $f(x_k)$ et en arrondissant le développement à l'ordre 2, on obtient :

$$f(x_k + p) \simeq f(x_k) + p^T \nabla f(x_k) + \frac{1}{2} p^T \nabla^2 f(x_k) p \stackrel{\text{not.}}{=} m_k(x_k + p) .$$

Avec le gradient et la matrice hessienne comme donnés aux définitions 2.16 et 2.17 de la section 2.4. On a donc :

$$\begin{aligned} m_k &= f(x_k) + \begin{pmatrix} p_1 & p_2 & \dots & p_n \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f}{\partial x_{k_1}} \\ \frac{\partial f}{\partial x_{k_2}} \\ \vdots \\ \frac{\partial f}{\partial x_{k_n}} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} p_1 & p_2 & \dots & p_n \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial^2 f}{\partial x_{k_1}^2} & \frac{\partial^2 f}{\partial x_{k_1} \partial x_{k_2}} & \dots & \frac{\partial^2 f}{\partial x_{k_1} \partial x_{k_n}} \\ \frac{\partial^2 f}{\partial x_{k_2} \partial x_{k_1}} & \frac{\partial^2 f}{\partial x_{k_2}^2} & \dots & \frac{\partial^2 f}{\partial x_{k_2} \partial x_{k_n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_{k_n} \partial x_{k_1}} & \frac{\partial^2 f}{\partial x_{k_n} \partial x_{k_2}} & \dots & \frac{\partial^2 f}{\partial x_{k_n}^2} \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \\ &= f(x_k) + \sum_{i=1}^n \left(p_i \cdot \frac{\partial f}{\partial x_{k_i}} \right) + \frac{1}{2} \left(\sum_{i=1}^n \left(p_i \cdot \frac{\partial^2 f}{\partial x_{k_i} \partial x_{k_1}} \right) \quad \sum_{i=1}^n \left(p_i \cdot \frac{\partial^2 f}{\partial x_{k_i} \partial x_{k_2}} \right) \quad \dots \quad \sum_{i=1}^n \left(p_i \cdot \frac{\partial^2 f}{\partial x_{k_i} \partial x_{k_n}} \right) \right) \cdot \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \\ &= f(x_k) + \sum_{i=1}^n \left(p_i \cdot \frac{\partial f}{\partial x_{k_i}} \right) + \frac{1}{2} \sum_{j=1}^n \left(p_j \cdot \sum_{i=1}^n \left(p_i \cdot \frac{\partial^2 f}{\partial x_{k_i} \partial x_{k_j}} \right) \right) . \end{aligned}$$

Ainsi, comme nous l'avons vu dans la section 2.5 du chapitre 2, pour trouver le vecteur p_k qui minimise $f(x_k + p)$, il suffit d'annuler la dérivée partielle de m_k par rapport à p :

$$\begin{aligned} \frac{\partial m_k}{\partial p} = 0 &\Leftrightarrow \nabla f(x_k) + \nabla^2 f(x_k) p = 0 \\ &\Leftrightarrow \nabla^2 f(x_k) p = -\nabla f(x_k) \\ &\Leftrightarrow p_k = -\left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k) . \end{aligned}$$

Pour garantir que cette direction est une direction de descente, il faut que la matrice hessienne $\nabla^2 f(x_k)$ soit définie positive (voir définition 2.9 de la section 2.2). Malheureusement, le calcul de cette matrice est bien souvent trop coûteux et il devient nécessaire de passer à une méthode dite *quasi-Newton* pour laquelle la matrice hessienne est remplacée par une matrice B_k symétrique définie positive dans l'expression de p_k .

Un *problème aux moindres carrés linéaire* est un problème d'optimisation sans contrainte du type :

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} \|Ax - b\|^2 ,$$

où $A_{(m \times n)}$ est une matrice réelle, $x \in \mathbb{R}^n$ et $b \in \mathbb{R}^m$. Dans le chapitre 2 (section 2.6), nous avons appliqué une méthode aux moindres carrés à une série statistique afin de représenter au mieux des couples de points par une droite de régression linéaire. Dans le cas d'un problème d'optimisation comme celui-ci, la méthode consiste à minimiser le carré des écarts entre la valeur réelle et la valeur approchée. De cette manière, nous pouvons trouver un vecteur $x^* \in \mathbb{R}^n$ qui soit la meilleure approximation, au sens des moindres carrés, du système linéaire suivant :

$$Ax = b .$$

Dans ce contexte, nous obtenons le gradient et la matrice hessienne de la fonction f comme suit :

$$\nabla f(x) = A^T(Ax - b) \text{ et } \nabla^2 f(x) = A^T A .$$

Puisque f est convexe, le point x^* tel que $\nabla f(x^*) = 0$ est un minimum global de f (voir théorème 2.2 de la section 2.5). On a, pour cela, les *équations normales* :

$$A^T Ax = A^T b ,$$

où $A^T A$ est semi-définie positive. Un tel système d'équations se résout via le calcul numérique par des algorithmes adéquats. Par exemple, la méthode du *gradient conjugué* est une méthode dont le pas de descente vaut :

$$\alpha_k = \frac{-(r_k)^T p_k}{(p_k)^T A p_k} ,$$

où $r_k = Ax_k - b$ et où les directions de descente sont *conjuguées par rapport à A*, c'est-à-dire :

$$(p_i)^T A p_j = 0 \quad \forall i \neq j .$$

On a l'algorithme suivant :

```

k=0
r0 = A*x0-b
p0 = -r0
faire tant que rk != 0
  alphak = -(rk)T*pk / (pk)T*A*pk
  xk+1 = xk + alphak * pk
  rk+1 = A * xk+1 - b
  betak = (rk+1)T*A*pk / (pk)T*A*pk
  pk+1 = -rk+1 + betak+1 * pk
k → k+1

```

Considérons à présent un *problème aux moindres carrés non linéaire* :

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} \sum_{j=1}^m r_j^2(x) ,$$

où $r_j : \mathbb{R}^n \rightarrow \mathbb{R}$, avec $j = 1, \dots, m$, sont appelés *résidus*. On définit alors le *vecteur résidu* $r : \mathbb{R}^n \rightarrow \mathbb{R}^m$ par :

$$r(x) = \begin{pmatrix} r_1(x) & r_2(x) & \dots & r_m(x) \end{pmatrix}^T .$$

De cette manière, la fonction à minimiser peut se réécrire, avec la norme euclidienne que nous avons définie dans le chapitre 2 (définition 2.12), comme :

$$f(x) = \frac{1}{2} \|r(x)\|_2^2 .$$

On définit la matrice *jacobienne* de $r(x)$ par :

$$J(x)_{(m \times n)} = \left(\frac{\partial r_j}{\partial x_i} \right) = \begin{pmatrix} \nabla r_1(x)^T \\ \nabla r_2(x)^T \\ \vdots \\ \nabla r_m(x)^T \end{pmatrix},$$

dont le calcul peut raisonnablement se faire sur ordinateur. Ceci nous donne :

$$\begin{cases} \nabla f(x) &= \sum_{j=1}^m r_j(x) \nabla r_j(x) = J(x)^T r(x) \\ \nabla^2 f(x) &= \sum_{j=1}^m \nabla r_j(x) \nabla r_j(x)^T + \sum_{j=1}^m \nabla r_j(x) \nabla^2 r_j(x) = J(x)^T J(x) + \sum_{j=1}^m \nabla r_j(x) \nabla^2 r_j(x) \end{cases}.$$

Ainsi, on obtient facilement le gradient $\nabla f(x)$ et le premier terme de la matrice hessienne $\nabla^2 f(x)$. En revanche, le calcul de $\nabla^2 r_j(x)$ ($j = 1, \dots, m$) est relativement coûteux mais, en général, les valeurs de cette matrice sont relativement petites et celle-ci est donc négligeable. On utilise donc l'approximation suivante :

$$\nabla^2 f(x) \simeq J(x)^T J(x),$$

et on retombe sur le cas linéaire où $r(x) = Ax - b$.

Dans une méthode quasi-Newton, on peut alors se servir du résultat ci-dessus en utilisant :

$$B_k = J_k^T J_k.$$

Chapitre 4

Le principe de prévision numérique du temps

Les prévisions numériques du temps sont, nous l'avons vu dans le chapitre 1, en constante évolution depuis plusieurs dizaines d'années. Dans ce chapitre, nous en expliquons le principe. Les informations ci-dessous proviennent des livres de NOCEDAL J. et WRIGHT S. J. [14] et de ASCHER U.M. et GRIEF C. [10], de la thèse de LALOYAUX P. [12], des articles de COIFFIER J. [19], de BLAYO E. et NODET M. [17] et de EMME-NEGGER F. [21], ainsi que du glossaire de MÉTÉO FRANCE [24]. Les explications reprises dans ce chapitre sont volontairement simplifiées de sorte que les lecteurs puissent comprendre et qu'un fil conducteur soit conservé. Nous rappelons que le but de ce mémoire n'est pas de s'étendre sur les prévisions météorologiques mais plutôt de montrer, par un exemple, la place que prennent les mathématiques dans le monde actuel.

Les prévisions numériques du temps se réalisent en quatre étapes distinctes. Premièrement, l'espace est décomposé en différents morceaux, différentes sections. On appelle cela une *discrétisation*. Ensuite, les valeurs initiales des différentes variables (température, pression, etc.) pour chacune de ces sections sont obtenues via des observations. La troisième étape consiste à modéliser les équations régissant l'atmosphère de sorte à obtenir un algorithme adéquat. Finalement, afin de pouvoir donner de nouvelles valeurs pour chaque section, on a recours à l'*assimilation de données*.

Une fois les prévisions numériques obtenues, les prévisionnistes ont encore la tâche de modifier ces résultats afin de les rendre accessibles au plus grand nombre après avoir analysé et sélectionné la prévision la plus pertinente.

4.1 Discrétisation de l'atmosphère

Résoudre un problème de prévision météorologique sur chaque partie de la Terre étant impossible, on passe à un modèle en dimension finie : on considère un maillage de l'atmosphère comme représenté à la figure 4.1. On constate ainsi que la surface de la Terre est décomposée en une grille d'un nombre fini de mailles ; l'atmosphère étant décomposé de manière volumique.

Avec ce type de décomposition, on peut estimer le temps de calcul pour une prévision à échéance H de la manière suivante :

$$T = \frac{N_v N_c N_t}{R} ,$$

où :

- N_v est le nombre de variables (vent, température, humidité, pression, etc.) ;
- N_c est le nombre de calculs à effectuer pour chaque variable et pour un pas de temps Δt ;
- $N_t = \frac{H}{\Delta t}$ est le nombre de pas de temps nécessaire pour atteindre l'échéance H ;
- R est la vitesse de calcul de l'ordinateur.

Notons que la condition CFL (condition de Courant, Friedrichs et Lewy), qui est une condition nécessaire à la convergence lors de la résolution numérique¹, doit être respectée :

$$U \frac{\Delta t}{\Delta x} < C$$

où U est la vitesse de propagation des ondes les plus rapides décrites par les équations, C un nombre sans dimension dépendant de la géométrie du problème et de la numérisation choisie et Δx la taille d'une maille.

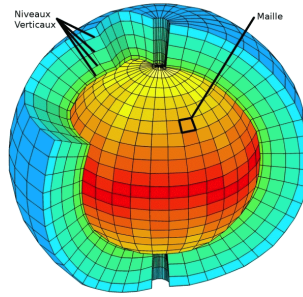


FIGURE 4.1 – Représentation du découpage en maillage. Source : Météo France.

Bien évidemment, plus le maillage est fin, plus les prévisions sont précises mais plus le temps d'exécution est élevé. Le défi est ainsi de trouver un bon compromis entre, d'une part, la précision des prévisions météorologiques et, d'autre part, un temps de calcul faible. Une piste de solution serait de passer à un modèle à mailles variables, utilisé dans certaines techniques de prévision. De cette manière, on peut obtenir une prévision plus précise sur un certain territoire et moins précise sur les alentours (voir figure 4.2).



FIGURE 4.2 – Représentation du découpage en mailles variables. Source : Météo France.

1. « Effet particulier de l'instabilité de calcul (erreur d'approximation dans des calculs numériques, qui grandit rapidement au fur et à mesure des calculs) associé à la prévision météorologique numérique. Si la dimension de la grille est inférieure à la distance parcourue dans l'intervalle de pas de temps par l'onde la plus rapide que permet l'équation, l'erreur grandit et envahit la solution physique. » (EUMETCAL [22])

4.2 Observations météorologiques

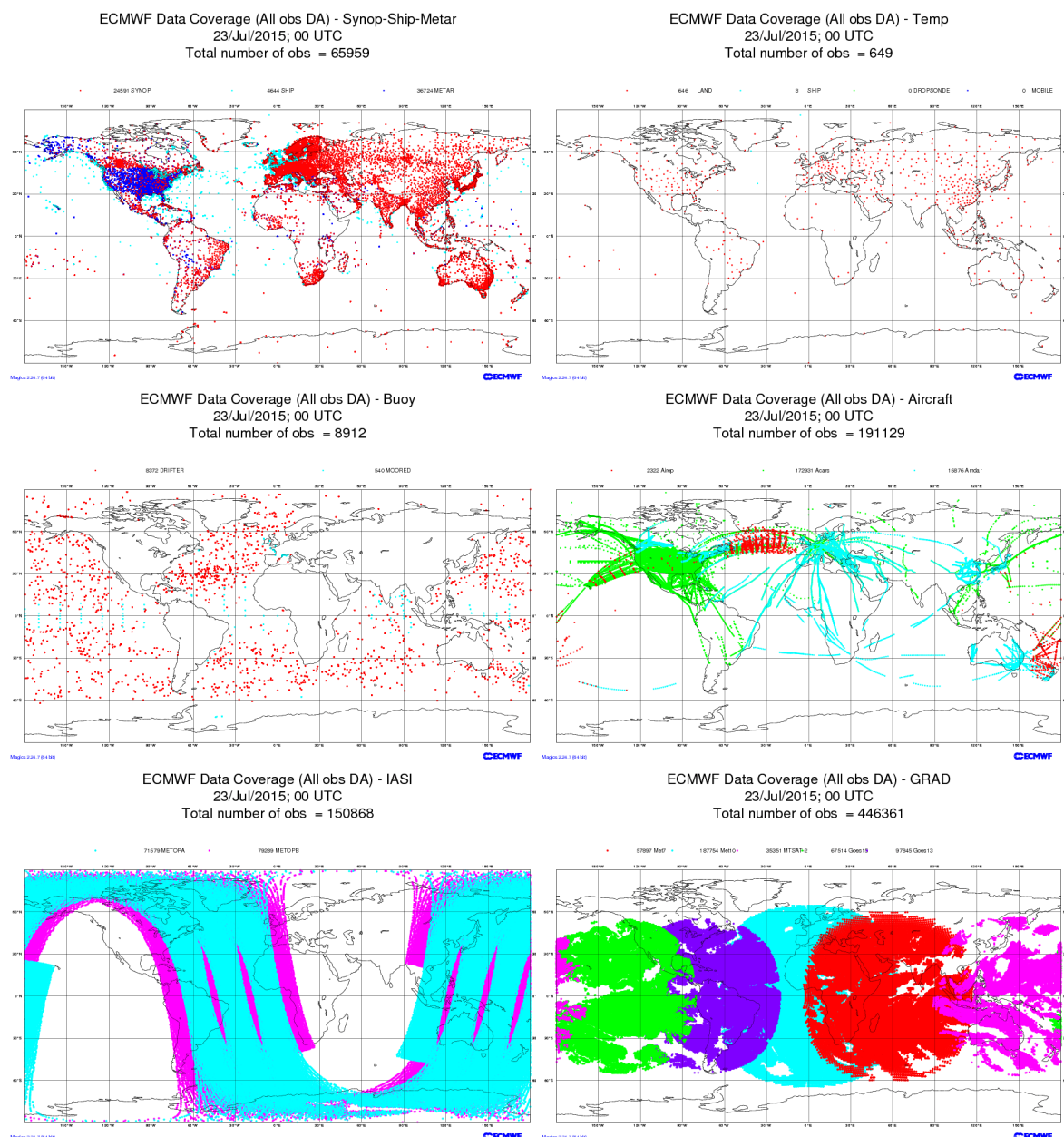


FIGURE 4.3 – Exemples d’observations obtenues ce 23 juillet 2015. Source : European Centre for Medium-range Weather Forecasts².

Les principaux moyens mis en place pour récolter toute une série de données d’observations sont les suivants (voir figure 4.3) :

- les stations météorologiques « en surface » : les stations terrestres (dont les informations sont envoyées en code Synop ou Metar) et les stations sur navire (Ship) qui mesurent la température, l’humidité de l’air, la pression atmosphérique, les précipitations, la force et la direction du vent ainsi que la durée d’ensoleillement ;

2. Le Centre Européen de Prévision Météorologique à Moyen Terme (CEPMMT ou ECMWF en anglais) est situé à Reading, en Angleterre, regroupant 18 pays européens, dont la Belgique. Il collecte toutes les données météorologiques nécessaires aux prévisions et développe des méthodes numériques.

- les stations aérologiques (radio-sondes suspendues à des ballons - Temp) qui permettent de mesurer la vitesse et la direction du vent ;
- les bouées océanographiques (Buoy) qui donnent la température de l'air et de l'eau des océans, la pression atmosphérique, l'humidité de l'air et la vitesse et la direction du vent ;
- les avions météorologiques (Aircraft) qui mesurent la température de l'air, la pression atmosphérique, la position des nuages, etc. ;
- les satellites (IASI et GRAD) qui envoient des images de la Terre ainsi que des observations concernant les nuages, la pression atmosphérique, etc.

4.3 Modélisation numérique

L'atmosphère est un fluide dont les phénomènes sont basés sur différentes lois de la physique : les lois de conservation de la quantité de mouvement, de la masse, de l'énergie, l'état de l'air, la loi de continuité, la loi de variation de la température (thermodynamique) ou encore la loi de variation de l'humidité.

Ces différentes lois sont traduites en équations par des personnes chargées de créer le modèle adéquat. La plupart des équations atmosphériques sont des équations aux dérivées partielles, que nous avons introduites dans le chapitre 2, section 2.4. Une fois le problème modélisé, les mathématiciens créent un algorithme dont se chargeront les ordinateurs. Ainsi, à partir de certaines conditions initiales et de conditions limites du domaine étudié, ceux-ci nous permettront d'obtenir des informations quant à l'état de l'atmosphère en un temps donné.

Prenons l'exemple de l'équation d'état de l'air. Cette équation est facilement compréhensible car elle ne fait intervenir aucune dérivée partielle. Celle-ci lie la pression, la température et la masse volumique de l'air. On considère que l'air de l'atmosphère est constitué d'air sec et d'eau à l'état liquide, solide et gazeux. On suppose aussi que l'air est un gaz parfait, c'est-à-dire qui satisfait toutes les lois des gaz et pouvant être décrit via la théorie cinétique des gaz. On a alors :

$$P = R\rho T$$

où P est la pression atmosphérique, ρ est la densité de l'air, R est une constante des gaz parfaits et T est la température. Ceci traduit le fait que si l'on maintient constante la pression de l'air, alors la densité et la température de l'air sont proportionnelles.

4.4 Assimilation de données variationnelle

L'assimilation des données est une technique d'estimation d'état initial d'un système (qui correspond à l'évolution de l'atmosphère en météorologie) dont le principe consiste à combiner des modèles mathématiques obtenus précédemment avec diverses observations pour en déduire des nouvelles conditions initiales et un modèle adapté. En pratique, il s'agit d'intégrer toute une série d'informations obtenues via des observations réalisées en cours de calculs de prédiction du temps dans ce même calcul afin de l'ajuster et de le rendre plus proche de la réalité.

En pratique, il est impossible de recueillir toutes les observations au même moment ni de manière homogène. Ainsi, un logiciel capable de faire intervenir de nouvelles informations en cours de calcul est tout à fait indispensable. Ce type de logiciel demande une

capacité énorme de stockage et de calculs. À intervalle régulier de temps, les résultats obtenus par le logiciel sont comparés avec les données d'observation disponibles à cet instant et le modèle est alors ajusté.

Considérons le vecteur d'état de l'atmosphère $x = (x_1 \ x_2 \ \dots \ x_n)^T$ et le vecteur des observations $y = (y_1 \ y_2 \ \dots \ y_p)^T$.

Par convention, on note x^b le vecteur (appelé *ébauche*) provenant de la précédente prévision de x et x^a le vecteur d'état analysé, solution d'un problème d'optimisation. Nous définissons l'opérateur :

$$\begin{aligned} \mathcal{H} : \mathbb{R}^n &\rightarrow \mathbb{R}^p \\ x &\mapsto \mathcal{H}(x) \end{aligned} ,$$

qui permet d'identifier une correspondance entre un vecteur d'état et un vecteur d'observations. Ainsi, un vecteur d'observations peut être lié au vecteur d'état réel (inconnu) du système par la relation suivante :

$$y = \mathcal{H}(x) + e^o ,$$

où $e^o \in \mathbb{R}^p$, appelé *erreur d'observations*, représente les erreurs dues aux instruments de mesures ainsi que les erreurs d'approximation lors de l'encodage des données. Si l'on considère que le modèle est linéaire, l'opérateur peut s'écrire sous la forme d'une matrice $H_{(p \times n)}$ et on a $y = H \cdot x + e^o$.

Soient :

- $B_{(n \times n)}$, la matrice de covariance d'erreurs d'ébauche $e^b = x - x^b$;
- $R_{(p \times p)}$, la matrice de covariance d'erreurs d'observations $e^o = y - \mathcal{H}(x)$;
- $A_{(n \times n)}$, la matrice de covariance d'erreurs d'analyse $e^a = x - x^a$.

avec la matrice de covariance définie dans le chapitre 2 (définition 2.26).

On définit une *fonction coût* qui donne, d'une part, une mesure de distance entre l'état et l'ébauche et, d'autre part, une mesure de distance entre l'état et les observations :

$$\begin{aligned} J(x) &= J^b(x) + J^o(x) \\ &= \frac{1}{2}(x - x^b)^T B^{-1}(x - x^b) + \frac{1}{2}(\mathcal{H}(x) - y)^T R^{-1}(\mathcal{H}(x) - y) . \end{aligned}$$

Le but étant d'obtenir un modèle qui suit au mieux la réalité en s'adaptant aux observations, tout en restant proche de l'ébauche, on cherche la valeur de x^a qui minimise la fonction coût. Il s'agit d'un problème d'optimisation ou, plus précisément, d'un problème aux moindres carrés.

Supposons que l'on soit dans le cas linéaire. Le théorème suivant nous donne une solution analytique du problème.

Théorème 4.1. Soient $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $M_{(m \times n)}$ une matrice réelle de rang n et N une matrice réelle symétrique définie positive d'ordre m . On définit la norme suivante :

$$f(x) = \|Mx - b\|_N^2 \stackrel{\text{déf.}}{=} (Mx - b)^T N (Mx - b) .$$

Alors, la solution de :

$$\min_{x \in \mathbb{R}^n} f(x)$$

est :

$$x^* = (M^T N M)^{-1} M^T N b .$$

Démonstration.

$$\begin{aligned} f(x) &= (Mx - b)^T N (Mx - b) \\ &= (Mx - b)^T (N^{\frac{1}{2}})^T N^{\frac{1}{2}} (Mx - b) \\ &= \left(N^{\frac{1}{2}} (Mx - b) \right)^T \left(N^{\frac{1}{2}} (Mx - b) \right) \\ &= \left((N^{\frac{1}{2}} M)x - (N^{\frac{1}{2}} b) \right)^T \left((N^{\frac{1}{2}} M)x - (N^{\frac{1}{2}} b) \right) \end{aligned}$$

On pose $A = N^{\frac{1}{2}} M$ et $c = N^{\frac{1}{2}} b$.

$$\begin{aligned} &= (Ax - c)^T (Ax - c) \\ &= x^T A^T A x - c^T A x - x^T A^T c + c^T c \\ &= x^T A^T A x - 2c^T A x + c^T c , \end{aligned}$$

d'où :

$$\nabla f(x) = 2A^T A x - 2A^T c = 2(A^T A x - A^T c) .$$

Ainsi, x^* est le minimum de $f(x)$ si

$$\begin{aligned} \nabla f(x^*) = 0 &\Leftrightarrow A^T A x^* = A^T c \\ &\Leftrightarrow x^* = (A^T A)^{-1} A^T c \\ &\Leftrightarrow x^* = \left((N^{\frac{1}{2}} M)^T (N^{\frac{1}{2}} M) \right)^{-1} (N^{\frac{1}{2}} M)^T (N^{\frac{1}{2}} b) \\ &\Leftrightarrow x^* = \left(M^T (N^{\frac{1}{2}})^T N^{\frac{1}{2}} M \right)^{-1} M^T (N^{\frac{1}{2}})^T N^{\frac{1}{2}} b \\ &\Leftrightarrow x^* = (M^T N M)^{-1} M^T N b \end{aligned}$$

■

Revenons à notre $J(x) = J^b(x) + J^o(x)$. On pose :

$$b = \begin{pmatrix} x^b \\ y \end{pmatrix} ; M = \begin{pmatrix} I \\ H \end{pmatrix} \text{ et } N = \begin{pmatrix} B^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix}$$

et, de cette manière, on a :

$$\|Mx - b\|_N^2 = \left(\begin{pmatrix} I \\ H \end{pmatrix} x - \begin{pmatrix} x^b \\ y \end{pmatrix} \right)^T \begin{pmatrix} B^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix} \left(\begin{pmatrix} I \\ H \end{pmatrix} x - \begin{pmatrix} x^b \\ y \end{pmatrix} \right)$$

$$\begin{aligned}
&= \begin{pmatrix} x - x^b \\ Hx - y \end{pmatrix}^T \begin{pmatrix} B^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} x - x^b \\ Hx - y \end{pmatrix} \\
&= \begin{pmatrix} (x - x^b)^T B^{-1} & (Hx - y)^T R^{-1} \end{pmatrix} \begin{pmatrix} x - x^b \\ Hx - y \end{pmatrix} \\
&= (x - x^b)^T B^{-1} (x - x^b) + (Hx - y)^T R^{-1} (Hx - y) \\
&= 2J(x) .
\end{aligned}$$

Puisque l'on cherche à minimiser cette fonction de x , on applique le théorème avec :

$$\begin{cases} M^T N M &= \begin{pmatrix} I^T & H^T \end{pmatrix} \begin{pmatrix} B^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} I \\ H \end{pmatrix} = \begin{pmatrix} B^{-1} & H^T R^{-1} \end{pmatrix} \begin{pmatrix} I \\ H \end{pmatrix} = B^{-1} + H^T R^{-1} H \\ M^T N b &= \begin{pmatrix} I^T & H^T \end{pmatrix} \begin{pmatrix} B^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} x^b \\ y \end{pmatrix} = \begin{pmatrix} B^{-1} & H^T R^{-1} \end{pmatrix} \begin{pmatrix} x^b \\ y \end{pmatrix} = B^{-1} x^b + H^T R^{-1} y \end{cases}$$

et la solution du problème d'optimisation qui consiste à minimiser la fonction coût est :

$$\begin{aligned}
x^* &= (M^T N M)^{-1} M^T N b \\
&= (B^{-1} x^b + H^T R^{-1} H)^{-1} (B^{-1} x^b + H^T R^{-1} y) \\
&= (B^{-1} x^b + H^T R^{-1} H)^{-1} (B^{-1} x^b + H^T R^{-1} x^b - H^T R^{-1} x^b + H^T R^{-1} y) \\
&= (B^{-1} x^b + H^T R^{-1} H)^{-1} ((B^{-1} + H^T R^{-1} H) x^b + H^T R^{-1} (-H x^b + y)) \\
&= x^b + (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} (y - H x^b) \\
&= x^b + K(y - H x^b)
\end{aligned}$$

où $K = (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}$ est appelée *matrice de gain*. En pratique, le calcul d'une telle matrice est impossible au vu des grandes dimensions des matrices B , R et H . Il faut alors utiliser une méthode d'optimisation par approche itérative pour trouver une approximation du gradient de J suffisamment proche de 0 :

```

Initialisation: x=x_0 et n=0
Faire tant que ||∇J|| > ε ou n>n_max:
| Calculer ∇J
| Mettre à jour x
| n → n+1

```

Le calcul du gradient de J se fait via une méthode, dite *méthode adjointe*, dont nous ne parlerons pas dans le cadre de ce mémoire, et la mise à jour de x se fait via une méthode de descente comme, par exemple, la méthode de Newton (voir chapitre 3, section 3.4).

Le modèle décrit ci-dessous s'appelle *3D-Var* puisqu'il s'agit d'une méthode variationnelle faisant intervenir des variables spatiales uniquement. Si on désire intégrer une variable temporelle, on passe alors à un modèle quadridimensionnel, ou *4D-Var*. Pour

cela, on définit un modèle :

$$\begin{aligned} \mathcal{M}_{i,i+1} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x_i &\mapsto \mathcal{M}_{i,i+1}(x_i) = x_{i+1} \end{aligned} ,$$

qui fait évoluer l'état du système au temps t_i vers celui au temps t_{i+1} . Si l'on considère que le modèle est linéaire, on note l'opérateur $M_{i,i+1}$ et il s'agit d'une matrice carrée d'ordre n . Dans ce cas, la fonction coût devient :

$$\begin{aligned} J(x) &= J^b(x) + J^o(x) \\ &= \sum_{i=1}^n (x - x^b)^T B^{-1} (x - x^b) + (Hx_i - y_i)^T R_i^{-1} (Hx_i - y_i) \end{aligned}$$

où $x_i = M_{0,1}M_{1,2}\dots M_{i-2,i-1}M_{i-1,i}x \stackrel{\text{not.}}{=} M_{0 \rightarrow i}(x)$ est l'état du modèle au temps i et où y_i est l'observation au temps i . Ce modèle *4D-Var* est actuellement utilisé dans de nombreuses stations de prévisions météorologiques.

4.5 Liens avec les notions vues en secondaire

Nous terminons ce mémoire en reprenant les grandes lignes du programme scolaire développé au point 2.1 du chapitre 2 afin de faire le lien entre les matières vues en secondaire et les prévisions météorologiques.

Géométrie

La géométrie vue en cinquième année approfondit la représentation dans l'espace étudiée en quatrième année. À l'évidence, la construction des modèles numériques, ainsi que la collecte de données atmosphériques, dépendent de la structure sphérique de la Terre. De plus, la discrétisation de l'espace ne pourrait se faire de manière efficace sans connaissance préalable de géométrie.

Algèbre

En algèbre, les élèves de cinquième année apprennent à modéliser une situation sous forme d'une matrice. Ils apprennent à effectuer différents calculs sur ces matrices et étudient des systèmes d'équations linéaires. Le principe de modélisation mathématique, enseigné tout au long des études secondaires, est à la base de milliers de problèmes de grande envergure. Ainsi, en météorologie, on représente les données recueillies par observations de tout type (radiosonde, satellite, etc.) sous forme de vecteur. On construit un modèle qui nous donnera, pour un temps donné t , un nouveau vecteur contenant les informations météorologiques (comme, par exemple, la vitesse du vent, la pression atmosphérique ou encore la température) à ce moment t . Cette étape cruciale se fait par l'intermédiaire de toute une série d'opérations matricielles.

Analyse

En analyse, les élèves de cinquième année apprennent, entre autres, à calculer la dérivée d'une fonction et à utiliser les propriétés des dérivées dans diverses applications. Cette

étude de dérivées s'étend aux notions de primitive et intégrale en sixième année. Nous avons vu comment généraliser la notion de dérivée à celle de dérivée partielle. Ainsi, les équations aux dérivées partielles sont à la base des prévisions météorologiques, pour modéliser la situation, représenter sous forme mathématique les lois qui régissent l'atmosphère, afin de pouvoir en donner l'état en un lieu et un moment donné. Par ailleurs, les dérivées de fonctions permettent une étude de celles-ci, décrivent leur courbe et en donnent les extremums. Les problèmes d'optimisation sont des conséquences de ces résultats. En météorologie, le principe d'assimilation de données consiste à optimiser l'estimation préalablement conçue de sorte à minimiser l'écart entre cette estimation et les observations récoltées.

Statistiques

Les statistiques étudiées en sixième année permettent, entre autres, aux élèves d'ajuster linéairement un nuage de points, c'est-à-dire de rechercher une droite qui représente le mieux possible la structure du nuage de points. Ils utilisent, pour cela, la méthode des moindres carrés. Afin de représenter la similitude entre cette étude et le principe de prévision météorologique, nous illustrons, d'une part, la régression linéaire et, d'autre part, le schéma d'assimilation de données.

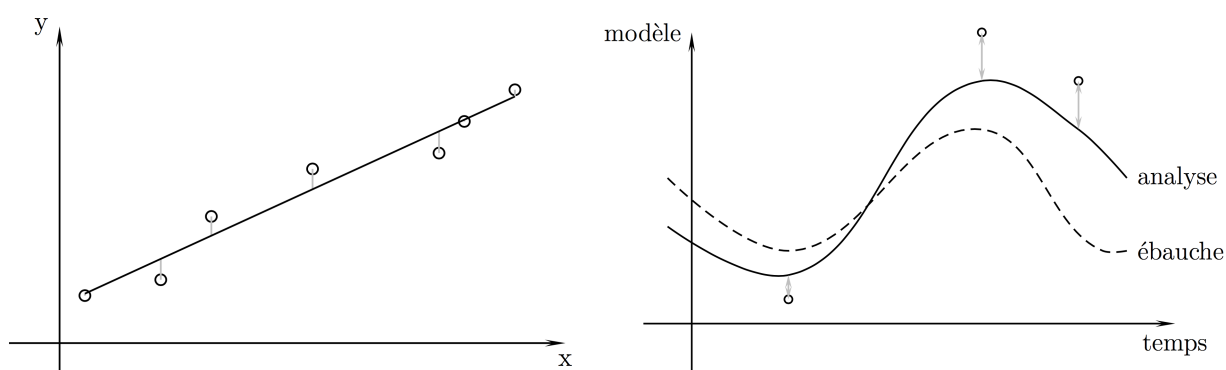


FIGURE 4.4 – Représentation, à gauche, d'une régression linéaire et, à droite, d'un schéma d'assimilation de données *4D-Var*.

On constate, sur le graphique de droite de la figure 4.4, que l'on a, en trait pointillé, l'ébauche, c'est-à-dire la précédente prévision du temps tandis que la courbe en trait plein est construite sur base de l'ébauche et est ajustée par rapport aux observations, représentées par des points sur le graphique.

Conclusion

Ce mémoire est divisé en quatre parties distinctes : une introduction historique des prévisions météorologiques, une étude du programme de mathématique au dernier cycle du secondaire, une synthèse du principe du calcul numérique et, finalement, une illustration de résolution de problèmes mathématiques, à savoir les prévisions numériques du temps.

Nous avons tout d'abord retracé l'évolution des prévisions météorologiques au cours du temps. Nous avons pu constater que la précision de celles-ci n'a cessé d'augmenter depuis plusieurs dizaines d'années. Cependant, l'atmosphère étant chaotique, c'est-à-dire qu'une petite variation initiale peut engendrer une modification significative dans le futur, il n'est pas imaginable de concevoir, un jour, des prévisions très précises et à très long terme.

Ensuite, nous avons analysé le programme de mathématique en cinquième et sixième secondaire et y avons repéré et développé les points en lien avec le calcul numérique : le calcul matriciel et les systèmes d'équations linéaires, les dérivées de fonctions (que nous avons pu étendre à la notion de dérivée partielle d'une fonction et à celle de problèmes d'optimisation sans contrainte) ou encore les statistiques et, plus particulièrement, la méthode des moindres carrés.

Nous avons, par la suite, détaillé les divers concepts du calcul numérique et avons ainsi pu établir des algorithmes généralisant certaines notions enseignées en secondaire comme la résolution de systèmes d'équations ou la méthode des moindres carrés.

Finalement, nous avons passé en revue les étapes de la prévision numérique du temps. Dans un premier temps, l'atmosphère est décomposée en une grille de points pour réduire l'espace de dimension infinie à un espace de dimension finie. Ensuite, diverses méthodes d'observation de données atmosphériques sont mises en place afin de récolter toute l'information nécessaire à la prévision du temps. Pour que cette prévision puisse être effectuée, une modélisation de l'atmosphère, à partir de systèmes d'équations différentielles principalement, est nécessaire : il s'agit de la troisième étape. Finalement, une fois le modèle formulé et les données récoltées, l'assimilation de données permet d'évaluer l'état de l'atmosphère dans le futur. Les calculs effectués à cette étape ne peuvent se faire autrement que via les ordinateurs et par méthode itérative étant donné la quantité d'informations qu'ils contiennent.

En guise de conclusion de ce mémoire, nous aimerions insister sur l'importance des mathématiques dans diverses applications concrètes : d'une part, il est indispensable d'avoir recourt aux ordinateurs qui effectuent des calculs bien plus rapidement que l'homme, et, d'autre part, une bonne connaissance des résolutions « manuelles » est primordiale dans la conception d'algorithmes efficaces.

Bibliographie

Cours

- [1] FÜZFA A., *Cours d'Introduction aux algorithmes mathématiques et au calcul scientifique*, Université de Namur, Année académique 2012-2013.
- [2] HARDY A., *Cours de Régressions linéaire et non linéaire*, Université de Namur, Année académique 2012-2013.
- [3] LEMAÎTRE A., *Cours de Mécanique des fluides*, Université de Namur, Année académique 2012-2013.
- [4] SARTENAER A., *Cours d'Algèbre linéaire numérique*, Université de Namur, Année académique 2013-2014.
- [5] SARTENAER A., *Cours d'Analyse numérique*, Université de Namur, Année académique 2012-2013.
- [6] STRODIOT J.-J., *Cours d'Introduction à l'optimisation*, Université de Namur, Année académique 2012-2013.

Livres

- [7] ADAM A., LOUSBERG F., *Espace Math 5^e/6^e : Tome 2*, de boeck, 2012, pp. 71-91.
- [8] ANNOYE M., GILON J.-L., VAN EERDENBRUGGHE A., WILEMME J., *CQDF Math 6^e, 6 périodes/semaine*, de boeck, 2014, pp. 96-104 ; 301-306 ; 322 ; 359.
- [9] ANNOYE M., VAN EERDENBRUGGHE A., *CQDF Math 5^e, 6 périodes/semaine*, de boeck, 2013, pp. 150-155 ; 204-219 ; 392-407.
- [10] ASCHER U.M., GRIEF C., *A first course in numerical methods*, Siam, 2011, pp. 17-31 ; 258-270.
- [11] DODGE Y., *Optimisation appliquée*, Springer, 2005, pp. 20-26.
- [12] LALOYAUX P., *Recent developments in optimization methods for data assimilation in oceanography*, FUNDP, 2012, pp. 7-11 ; 17-24 ; 61-67.
- [13] LAY D. C., *Algèbre linéaire et applications*, 4^e édition, Pearson France, 2012, pp. 101-114 ; 126-128 ; 176-178.
- [14] NOCEDAL J., WRIGHT S. J., *Numerical Optimization*, Springer, 1999, pp. 11-32 ; 35-55 ; 252-259.
- [15] QUARTERONI A., SACCO R., SALERI F., *Méthodes Numériques : Algorithmes, analyse et applications*, Springer, 2007, pp. 5-9 ; 27-28 ; 47-53.

Documents électroniques

- [16] ANDERSON E., *World Weather Watch successes and challenges for the future*, disponible sur <http://www.wmo.int/pages/publications/bulletinarchive>, consulté le 20/07/2015.
- [17] BLAYO E., NODET M., *Introduction à l'assimilation de données variationnelle*, disponible sur <https://team.inria.fr/moise/files/2012/03/Methodes-Inverses-Var-M2-math-2009.pdf>, consulté le 26/07/2015.
- [18] BOUQUET A., *Le programme Manhattan*, disponible sur <http://evelyne.bouquet.free.fr/WebAlain>, consulté le 20/07/2015.
- [19] COIFFIER J., MÉTÉO FRANCE, *Un demi-siècle de prévision numérique du temps*, disponible sur http://documents.irevues.inist.fr/bitstream/handle/2042/36122/meteo_2000_30_1, consulté le 20/07/2015.
- [20] EC JOURNAL 2008, *Sprott/Lorenz*, disponible sur <http://sprott.physics.wisc.edu/lorenz.pdf>, consulté le 23/07/2015.
- [21] EMMENEGGER F., *Matières et chimie*, disponible sur <http://chimge.unil.ch/Fr/mat/1mat7.htm>, consulté le 09/08/2014.
- [22] EUMETCAL, *Instabilité de Courant, Friedrichs et Lewy*, disponible sur <http://www.eumetcal.org/euromet/glossary/cflcoura.htm>, consulté le 13/08/2015.
- [23] FÉDÉRATION DE L'ENSEIGNEMENT SECONDAIRE CATHOLIQUE, *Programme : Mathématiques - 3^e degré - Humanités générales et technologiques*, disponible sur <http://admin.segec.be/documents/4473.pdf>, consulté le 09/04/2014.
- [24] MÉTÉO FRANCE, *Glossaire*, disponible sur <http://www.meteofrance.fr/publications/glossaire>, consulté le 26/07/2015.
- [25] MINISTÈRE DE LA COMMUNAUTÉ FRANÇAISE, *Enseignement Secondaire Ordinaire De Plein Exercice; Humanités Générales Et Technologiques; Enseignement Secondaire Général Et Technique De Transition - Troisième degré - Programme d'études du cours de : Mathématiques*, disponible sur <http://www.wallonie-bruxelles-enseignement.be/prog/40-2000-240.pdf>, consulté le 09/04/2014.